Holographic Optical Interconnects in Dichromated Gelatin

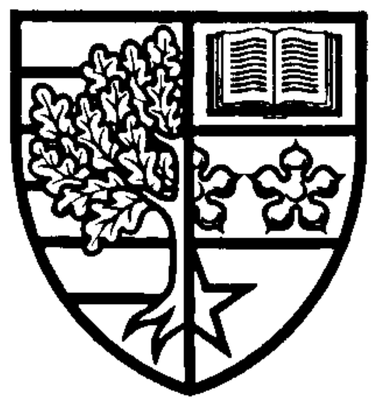
Edward John Restall

Submitted for the degree of Doctor of Philosophy.

Heriot-Watt University

Department of Physics

February 1997.



Contents

List of Figures

List of Tables

Acknowledgements

I would like to thank my supervisor Professor Andrew Walker for all his help, support and guidance over the past eight years. Thanks also to Professors Des Smith and Carl Pidgeon for their support and for giving me the opportunity of working in the physics department.

This thesis builds largely upon the previous work of Ian Redmond who has given me encouragement beyond the call of duty during my extended period of research - cheers Ian - it is only by standing upon his shoulders that I have managed to produce this. I would also like to thank Brian Robertson for his warm friendship and acknowledge his collaboration on a large part of the space-variant work contained in chapter 7. The optical vector-matrix crossbar and O-CLIP systems described in the course of chapter 6 were departmental projects: the bulk optics of these systems were designed, assembled and tested by Ian Redmond & Doug McKnight, the binary phase elements were designed and fabricated by Mike Miller - their work is duly acknowledged.

To all the members of the diffractive optics groups (past and present) at Heriot-Watt and Helsinki University of Technology I would like to say a big thank you, especially: Mike Miller, Ian Barton, Jari Turunen and Antti Vasara. Thanks also to Mo Taghizadeh for the use of his SUN system and to Paul Blair & Ian Galbraith for managing it. The inspiration for most of the theoretical work in this thesis were M. G. Moharam and T. K. Gaylord - though they weren't to know - thanks anyway. The creative stimulus for most of the experimental work contained in this thesis was driven by the rolling programme of optical computing research within the physics department - thanks to everyone involved in this.

This PhD has been supported by the Science and Engineering Research Council through a CASE studentship in collaboration with British Aerospace plc Sowerby Research Centre in Bristol - thanks also to the Engineering and Physical Sciences Research Council, even though they wrote me off as incomplete. I am grateful to Nick Roberts, Nigel Aldridge and Henry White at BAe for their support.

Thanks to Rebecca, Alan and Tony (my office mates for the first three years),

for their friendship along the learning curve that we followed together. Also to Frank Tooley (and family), Robert Craig, Jon Milward, John Simpson, Paul Barrett, Dominic Goodwill, Dave Neilson, Matt Derstine, Miles Murdocca, Tom Stone, Rick McCormick and Ashley Lloyd. To every member of the academic, secretarial and technical staff in the physics department who has shown me friendship and support over the last eight years I would like to express my sincere thanks.

Thank you to Professors Trevor Hall and Frank Tooley for so readily agreeing to be my external and internal examiners and for making the time to assess this thesis.

Finally, I would like to thank my family: Mum, Dad, Elizabeth, Maria, Richard and Claire, and close friends: Simon, Nicky, Sean, Carrie, Helen and Kathryn for their unstinting love and encouragement, without which this thesis would never have been completed.

Ed Restall,

Heriot-Watt University,

Edinburgh.

Abstract

The requirements of an ideal recording medium for volume holographic interconnects are considered and compared with the properties of dichromated gelatin (DCG). Practicalities of recording such holographic optical elements (HOEs) in DCG are described along with various theories advantageous to the design and fabrication of a range of transmission and reflection free-space interconnects. A rigorous coupled-wave theory (RCWT) is presented to model the performance of amplitude and phase modulated planar grating structures, including grating harmonics. This theory was applied to experimental data for both transmission and reflection gratings and a comparison made with multi-wave coupled-wave theory (for transmission). It is shown that the limited spectral sensitivity of DCG can be overcome, so that interconnects can be fabricated for replay in the near infra-red, (outside the recording sensitivity range). A theory is developed to incorporate a wavelength-shift between recording and replay for HOEs that contain focusing power. Using this theory wavelength-shifted holographic lenslets were produced that operate at or close to the diffraction limit for $f/3$ or greater and represent the best performance that can be achieved using spherical recording optics. The RCWT was extended to analyse fan-out holograms and applied to volume phase gratings. The theory has been found to quantitatively match published experimental data for multiplexed HOEs recorded both sequentially and simultaneously. A detailed theoretical analysis of reflection fan-out has shown that higher efficiencies and better uniformities than equivalent transmission gratings are achievable, e.g. for small fan-out angles (without grating phase control) the maximum useful diffraction efficiency of transmission gratings, as the level of fan-out increases, approaches 48%, compared with 78% for similar reflection structures. This theory represents the most generally applicable, accurate and numerically solvable model to date. Finally a selection of spatially invariant and variant interconnects are described. These were designed, using the theories developed in the thesis, for use in experimental optical communication and processing circuits and were fabricated using the DCG techniques described.

Preface

This thesis represents my three years of research as a PhD student between 1988 and 1991. Unfortunately, before I managed to write-up I was taken ill with what was later diagnosed as Myalgic Encephalomyelitis (ME); also variously known as chronic fatigue or post-viral fatigue syndrome. This is a little understood but chronic and disabling illness that restricts the sufferer both physically and mentally [COL94, FOX96], often for many years. As a consequence of this, most of the writing (and the modelling contained in chapter 5) has been carried out over the last two years whilst recovering from the illness.

The completion of this work could not have been achieved without the support of the university and in particular the physics department. I also owe a great debt of thanks to the school of computing and mathematics at the university of Teesside for providing me with the internet access, unix system facilities and PC networking expertise that has enabled me to write-up from my parents home. Thankyou especially to Tim Skipper, Jayne Kennedy, John Tunstall, Richard Finn (for his LaTeX help [LAM86]), Mark Payne, John Marsden and Dave Restall.

My supervisor, Andy Walker, has been particularly patient with me under the circumstances. He has made every effort to accommodate the restrictions imposed upon me by my illness, right down to the provision of a fluorescent lighting free working environment and most importantly of all the laptop PC on which the thesis was written. For all this and for his friendship - thanks again Andy.

I have been fortunate in having a G.P., Dr C. P. O'Neill, who has kept an open mind on my illness when many other medical doctors, professing to be men and women of science, close their minds to anything that they don't understand or haven't experienced. Thank you to him and also to Judy Sheehan and Wendy Francis for their expertise in helping me to adapt to, accept and cope with the dramatic change in lifestyle that my illness has caused.

As a direct consequence of the timing of this long illness, the vast majority of the work contained in this thesis remains unpublished in either journals or conference proceedings.

To Mum & Dad

Chapter 1

Introduction

Optical information processing systems began to be of interest in the 1960s as a way of performing complex *analogue* computations both quickly and in parallel. Signal processing systems, using coherent light, such as synthetic aperture radar [CUT66] and Vander Lugt correlators [VAN64], were based around the Fourier transforming capabilities of a lens system [MCA91, chapter 3]. Although *digital* optical processing and communication systems had been proposed, it was not until the advent of low power optical logic devices (optical analogues of the transistor) working at practical wavelengths that they began to be realisable.

The first demonstration digital optical computing system performed sequential binary logic operations on a 2D array or image [JEN84]. The optical logic gates were implemented using a liquid crystal light valve (LCLV) and computer generated holograms (CGHs) were used to provide gate-to-gate interconnections. Since then, the invention of optically bistable switching devices, that are faster and consume less power than the LCLV, has led to numerous demonstrations of computational circuits and photonic switching systems [WHE95]. This thesis concentrates upon optical systems based around two different devices:

- The optothermal non-linear interference filter (NLIF) which is based on the $II-VI$ compound zinc selenide (ZnSe), [ABR83, SMI87, WAL91]; and
- The optoelectronic self electro-optic effect device (SEED) that is made from compounds of the gallium aluminium arsenide (GaAlAs) $III-V$ semiconduc-

tor family, [MIL84, PRI90, TOO93, HIN94].

Information processing systems that use these devices manipulate 2D arrays of optical data channels (see section 1.2) in a similar way to the original LCLV system.

The growth of more sophisticated switching devices was accompanied by a parallel development of interconnection systems required to provide the "optical wiring" between data planes. The need for efficient, cheap and lightweight micro-optic connection systems has led to the field of research known as optical interconnection technology. This embraces bulk-optics, micro-optics, planar optics, waveguide and diffractive optics. The field has developed from the provision of communication paths for digital optical systems to the general study of optical interconnects at all levels between electronic, optoelectronic and all-optical information processing systems, e.g. [MCA91, WHE95].

This thesis concentrates on the modelling, design and optical recording of novel diffractive interconnects created in volume media. Particular emphasis is placed upon elements recorded optically in dichromated gelatin (DCG) although the models and discussion are applicable to other volume holographic media. DCG is an almost ideal material for the recording of holographic optical elements (HOEs): it can produce diffraction efficiencies approaching 100% (in reflection or transmission), can be used for operation in the visible and near infra-red (encompassing most current communication wavelengths) and can be processed to have a low absorption and a high signal-to-noise ratio (SNR). This research builds upon the routing, focusing and fan-out HOE functions explored by Redmond [RED89a] and was carried out in parallel with the free-space interconnection work of Robertson [ROB93a].

The rest of this chapter explains the motivation for implementing optical rather than electronic interconnects and presents an architectural model useful for discussing optical communication systems. Different classes of interconnect are described along with some design criteria that are important in the implementation of optical interconnection systems. Finally a description of the coordinate system used throughout and an outline of the thesis is given.

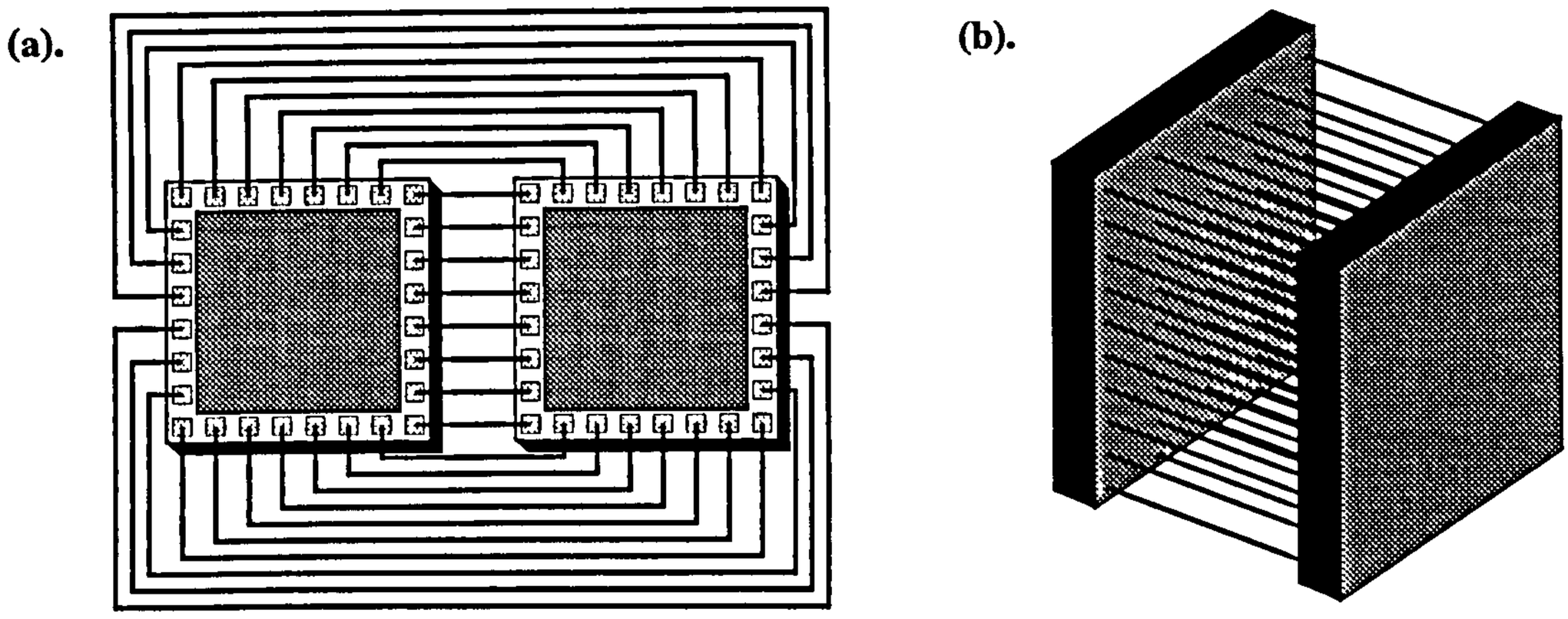


Figure 1.1: The difference between the interconnections (one-to-one) of electronic and optoelectronic chips: (a). Quasi-planar edge connections of electronics; and (b). Parallel surface connections of optics (after Murdocca [MUR88]).

1.1 Reasons for optical processing systems

Since the invention of the transistor, in the 1950s, significant advances have been made in microelectronic design and fabrication techniques. These have led to ever faster and more powerful computing and communication systems. However, VLSI technology has reached the point where limitations to data communication are beginning to restrict the performance of these systems. In addition, advances in multiprocessor and parallel computer architectures have further compounded the communications bottleneck. It is possible to take advantage of both the temporal (channel speed) and spatial bandwidth (channel connectivity) of optical and optoelectronic systems to overcome many of the communication limitations of electronics. The main areas where optics has an advantage over electronics in signal processing systems are:

- **Connectivity.** Photons, unlike electrons, do not readily interact with one another. Consequently light beams can pass through one another without distorting any information being carried. This allows for a higher packing density of communication channels (a well designed lens can transfer $\sim 10^6$ separate channels between data planes) and greater flexibility in the interconnection patterns between them. Optics allows 3D interconnections in free-space, which

offers significant advantages over the quasi-planar 2D interconnections of electronics (see figure 1.1).

- **Parallelism.** The 3D nature of optical interconnections permits the manipulation of 2D images or arrays of communication channels in parallel (see figure 1.2). This inherent parallelism of optical systems makes them a particularly attractive technology for parallel processing architectures. Conventional single instruction single data stream (SISD) computers suffer from a communications bottleneck brought about by the limited number of interconnections that can be supported by electronics. This is known as the *von Neumann bottleneck* [BAC82] and implies a performance limitation imposed by the sequential and address-oriented communications (along narrow bandwidth channels) between the central processing unit (CPU) and memory of a computer architecture based upon a modified finite state machine [HUA84]. The parallel communication capabilities of optical systems have the potential to overcome this sequential architectural limitation.

- **Pin-outs.** Although many thousands of components can be assembled on a single electronic integrated circuit (IC), a large proportion of the chip area can be devoted to wiring, (up to 70% in some cases). Optoelectronic chips allow a greater on-chip component count by utilising the 3D pin-out capability of optics with optical wiring off the surface of the IC, (see figure 1.1). In addition, a typical electronic IC has only a few hundred (edge-connected) pin-outs whereas optoelectronic chips have many thousands across the surface area. Flip-chip bonding [LEN96] allows silicon ICs to have an optoelectronic interface for 3D connectivity.

- **Transmission bandwidth, crosstalk and clock-skew.** Electronic interconnection technology is limited by several electro-magnetic (EM) conduction characteristics:

  - The skin effect [DUF80, page 243] leads to greater attenuation of signals in transmission lines at high frequencies.

- The characteristic impedance of transmission lines (e.g. a data bus) can cause signal distortion, due to reflections and ringing, if not correctly matched throughout the system. This is more problematic when source and/or load impedances vary with time or frequency; and
- Parasitic capacitance and inductance reduce the usable bandwidth of the interconnects, causing propagation delays. This is particularly problematic in the distribution of clock signals, because at faster clock rates any relative difference between the signals reaching different parts of the circuit will lead to non-synchronous operation (*clock-skew* [MEA80, page 230]). In addition to this inductive noise introduces EM crosstalk between neighbouring channels.
- The resistance-capacitance (RC) time-constant of interconnecting conductors limits the communication speed (data rate) between electronic switching elements.

These represents a fundamental limit to the performance of purely electronic systems. An estimate for the maximum *aggregate* bandwidth of a 2D electronic interconnection system has recently been made [MIL95, SMI95] as $B_{max} \sim 5 \times 10^{14}\ {}^{A}/_{D^2}$ Hz, where A is the total area of the electronic interconnections and D is the average length of each channel. For a 10cm connection, (e.g. across a circuit board), B_{max} is typically 150GHz. Optical systems have the potential to provide a much larger transmission bandwidth (> 100GHz per channel) because there is no distance dependent capacitative loading. Once in the optical domain signals can be sent any distance without a bandwidth penalty (ignoring transmission losses).

Optical systems have no physical contact between connections and no feedback to the power source (thereby avoiding data dependent loading). In addition, the inherently low dispersion of optical communication channels means that, even at high data rates, signal pulses retain their shape.

- **Energy efficiency.** Theoretical models for power consumption [FEL88] show that for communication distances of a few centimetres or less (i.e. across an IC) electrical connections are more efficient than optical alternatives. Over larger distances, however, optical interconnects consume less power and would consequently be more energy efficient for chip-to-chip, through to network level communications. In addition, high performance electronic processors consume a lot of electrical power to provide computing and communication speed, requiring processor elements to be packed close together in order to limit communication delays. This causes heat dissipation problems that are expensive to solve. The larger bandwidth and lower transmission loss of optical interconnects makes high performance computing possible with a physical architecture that is less tightly packed, thereby alleviating power consumption and heat dissipation costs.

- **Immunity to EM noise.** This is an area of increasing importance with European Union regulations requiring the shielding of all new communications equipment so that it does not cause interference with other systems. Unlike electronic communications, optical systems do not radiate over a range of frequencies. In addition, photons are immune to interference problems (see connectivity above) making optical systems more tolerant to external influences. Combined with this is the increased safety of optical systems which represent a lower ignition source risk in hazardous environments.

These main points represent a convincing argument for using optics at nearly all the levels of communication in information processing systems. There are the additional possibilities of: utilising the large bandwidth of optical systems to implement time-division multiplexing of many signals onto a single optical channel, wavelength-division multiplexing (not particularly suited to diffractive interconnects) that further increases transmission bandwidth and the use of optical storage media (e.g. photorefractives and optical discs) to provide a higher data storage density and access speeds than current magnetic and solid-state technology [RES89, MCA91].

1.1.1 Advanced computer architectures

Some effort has been made to implement optical versions of advanced sequential computer architectures, e.g. a reduced instruction set computer (RISC), [MCA91, chapter 13]. However, such architectures do not take full advantage of the massive parallelism of optical systems. Considerably more work has gone into exploiting the parallelism of optical interconnects for optical neural networks (as processors, adaptive learning systems or associative memory systems) [MCA91, chapter4 16 & 17], [WHE95, section IV]. The majority of the research into digital optical computers has concentrated on exploiting the parallelism of optics for single instruction multiple data-stream (SIMD) and multiple instruction multiple data-stream (MIMD) architectures:

Communication bottlenecks in conventional electronic systems [SIE82] have led to advanced computer architectural principles (non-Von Neumann models) such as *dataflow* [DEN80, GUR85] and *reduction* [BAC78, MAG80]. These decentralise processing and memory units within a computer, resulting in multi-processor parallel architectures [TIB84] capable of MIMD computing. The highly parallel nature of these electronic super-computer models has led to obvious limitations in their implementation that optical interconnections can circumvent. Optical dataflow architectures and their advantage over electronic equivalents are discussed in more detail by McAulay [MCA91, chapter 14].

Specialised digital computer architectures have been developed to take advantage of the parallelism and connectivity of optics [WHE95]. Since the first demonstrations of all-optical circuits [ABR83, WAL90] within the physics department at Heriot-Watt university, several SIMD optical processors have been designed and/or demonstrated. These were envisaged as subsystems to a host electronic computer, performing specialised parallel processing tasks, e.g. the cellular logic image processor (CLIP) [CRA91, WAL91] based upon NLIF switching devices (e.g. [BUL89]) and local interconnections (see section 6.2); a SEED (e.g. [LEN93]) version of the CLIP [TOO93]; and a perfect shuffling machine [DES95] based on SEEDs and non-local interconnects (see section 1.3.4).

Most of the research contained within this thesis was designed to provide interconnects for the optical computing architectures investigated within the department. However, they are equally useful in photonic switching fabrics.

1.1.2 Photonic switching architectures

Switching fabrics for local area networks (LANs), wide area networks (WANs) and telephony systems can be based upon space, frequency or time switch architectures. In addition to this, communication can be via packet (e.g. ethernet) or circuit switches (e.g. traditional telephone network) [DUN94, chapter 13]. Not all of these network architectures can be easily implemented using free-space photonics. However, considerable effort has been made in the area of free-space packet-switches [CLO93, HIN94]. In particular the class of interconnects known as extended generalised shuffle networks (EGSNs) represent a flexible architecture suitable for optical implementation [CLO92]. Such shuffle networks are based upon spatially variant interconnect (SVI) topologies (discussed in section 1.3.2 and chapter 7).

Packet switches are more bandwidth efficient than circuit switches and with the advent of asynchronous transfer mode (ATM) networks they are overtaking the use of circuit switched systems (even for real-time voice communications) [DUN94, chapter 10]. Consequently, photonic switching fabrics (and associated interconnects) based on packet switches are of increasing interest for both telecommunications and distributed computing systems.

1.2 Optical architecture model

A model for optical communication/processing systems is also shown in figure 1.2 to establish a standard set of terms for this thesis. The input and output planes can be referred to as the array or field. Each element in these planes is referred to as a pixel, communication node or channel - they can be physically spread-out in an "adhoc" manner, but are usually grouped together into a regular array. The optical interconnection system maps the input plane to the output plane, the mapping being defined by the interconnection pattern/topology under consideration. The size of

the input plane ($j \times k$) does not have to equal the output plane ($p \times q$), although they are usually equivalent. There can be several stages, l, to the optical interconnection system in order to achieve the connectivity pattern between pixels. Each stage can be pixelated into an array of $n_l \times m_l$ elements.

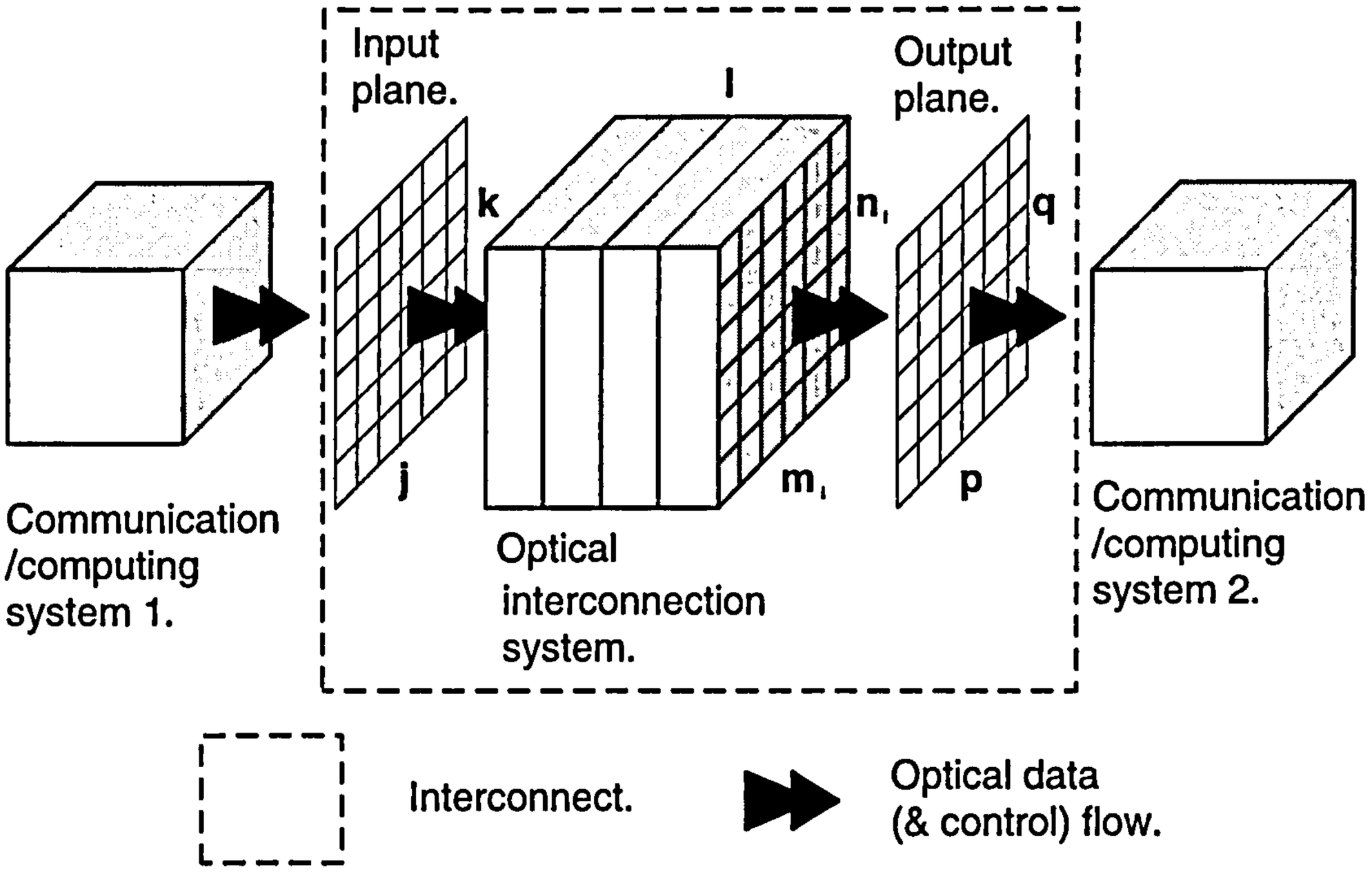


Figure 1.2: Standard architecture model used in this thesis for optical communication and computing systems, (the communicating/computing systems can be a collection of processors).

The model applies to all-optical or hybrid optical and electronic systems where communications are carried out using an optoelectronic interface at the input and output planes. The interconnection system can be static or dynamic and may possess some limited computational ability (e.g. smart-pixels [LEN93]) which may be controlled by systems 1, 2 or both. The processing systems 1 and 2 can be distinct and autonomous communicating units or sub-systems of a single unit. Alternatively they can represent the input and output communication ports of the same processing unit. Consequently, separate entities in the whole system (e.g. a processing unit in an MIMD computer) can communicate with all other entities in that system. The output can be fed back into the system to achieve more complicated interconnections, this feedback can be optical (within the interconnection block shown) or

electronic (via systems 1 and 2) depending upon the system under consideration. All of the interconnects discussed in chapters 6 and 7 of this thesis can be mapped onto this model and they are described in the terms outlined here.

1.3 Architectural considerations

An optical interconnection pattern can be implemented in a spatially variant or invariant manner, both of which can imply one-to-one connectivity or different levels of signal splitting and recombination. The relevance of this is outlined in this section together with an explanation of the considerable interest in shuffle networks.

1.3.1 Fan-in and fan-out

If an input data channel is optically connected to N output channels, that input channel is said to have N-fold fan-out. Conversely, if N channels in the input are connected to one channel in the output, that output channel is said to have N-fold fan-in. Chapter 5 and appendix C discuss optical fan-in and fan-out in more detail.

1.3.2 Space-variant & space-invariant optical interconnects

A strict definition of spatial invariance in a linear system states that: a shift in the position of the input merely changes the location of the output without altering its functional form. Such a definition can be used to ascertain the spatial variance or invariance of an optical imaging system. However, when discussing interconnects a slightly broader definition regarding the regularity of the connection pattern is more usual [JEN84, TOO94]: An interconnect is said to be space-invariant if the connection (fan-out) pattern is constant across the entire system (array).

A distinction is made between the connection pattern and the optical implementation of the interconnect. It is possible to have an optical system that has a separate interconnect hologram associated with each input pixel. If, however, the function carried out on each pixel in the input (array) is essentially the same (regardless of spatial location) the connection pattern is clearly space-invariant. It is equally

possible to have a space-invariant implementation of a space-variant interconnect, e.g. photonic switching systems demonstrated by AT&T in which shuffle equivalent interconnects are achieved using binary phase gratings [CLO94, HIN94].

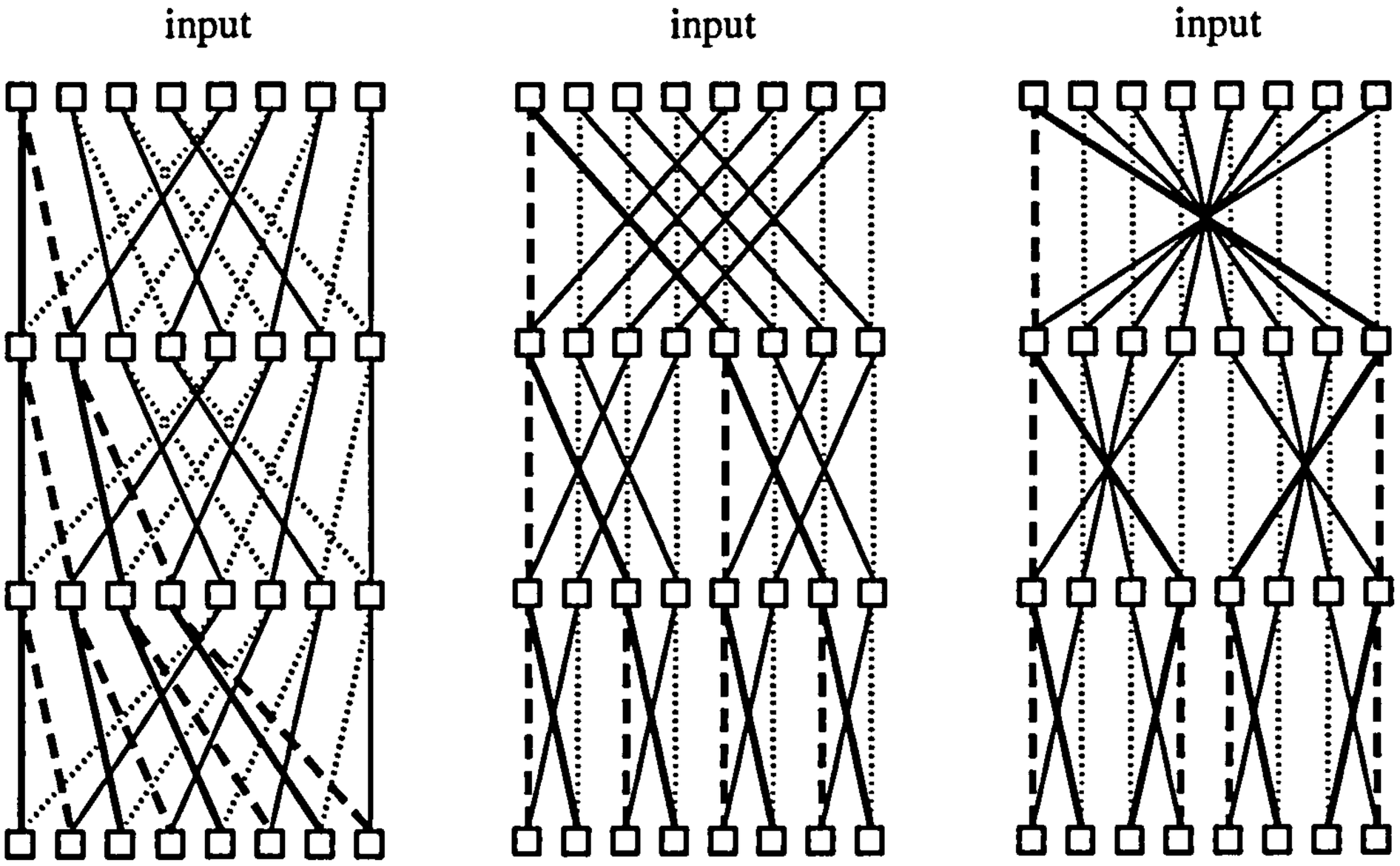


Figure 1.3: Examples of regular multi-stage space-variant interconnection networks: (a). the perfect shuffle, (b). the Banyan; and (c). the crossover. Examples are for logarithmic networks with 8 channels, requiring 3 stages for full interconnection; logic-planes/switching devices separate each interconnection stage, (heavier pathways illustrate how input node 1 is connected to *all* output nodes).

1.3.3 Shuffle networks

SVIs can be used to implement several multi-stage interconnection networks. Figure 1.3 illustrates the use of three SVI systems (the perfect shuffle, Banyan and crossover networks) to implement full connectivity between the input and output channels of the network. These are examples of logarithmic networks, in which alternating layers of optical processing elements (simple switching nodes or so called *smart pixels* [HIN94]) and interconnection stages provide the required routing and connectivity. Each interconnection stage is limited to a fan-in/fan-out of two and it takes $\log_2 N$ stages to fully interconnect N channels. The networks illustrated in fig-

ure 1.3 implement the same connectivity and are therefore topologically equivalent.

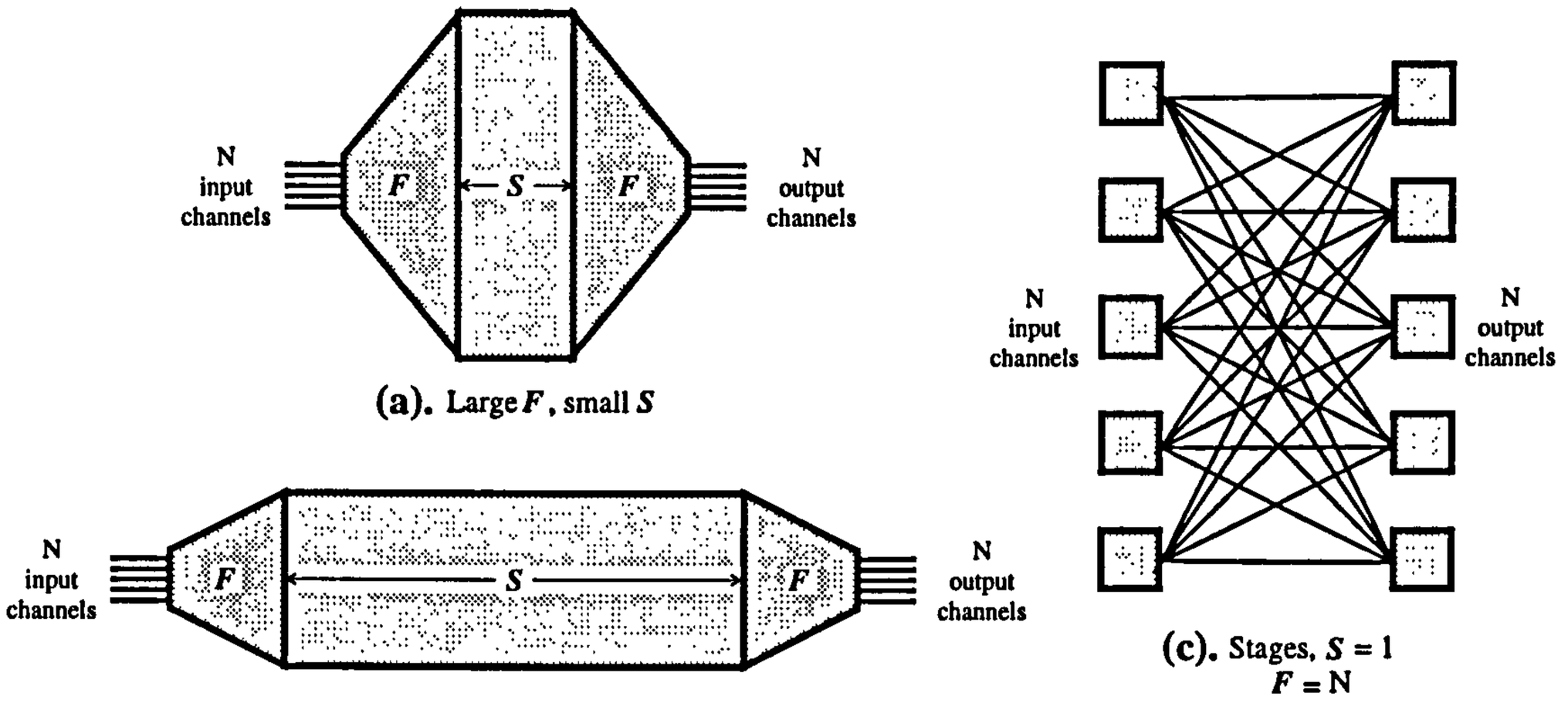


Figure 1.4: Extended generalised shuffle networks (after Cloonan et al. [CLO92]): (a). High level of fan-out/fan-in with few interconnection stages (short, fat network), (b). Low level of fan-out/fan-in with a lot of interconnection stages (long, thin network); and (c). The extreme of fan-out to N with 1 stage (crossbar network).

Logarithmic shuffle networks are a subset of the extended generalised shuffle networks (EGSNs). EGSNs utilise the same SVI patterns, however, fan-out of F is permitted between channels before the network followed by fan-in of F after the network and any number of stages, S, is allowed [CLO92]. The additional complexity of EGSNs is required to reduce the blocking probability to an acceptable level. This is particularly important for ATM packet switching systems. Cloonan et al. [CLO92] show that to achieve a desired transmission probability there is a trade-off between F and S (see figure 1.4a & b). The extreme (non-blocking) case is shown in figure 1.4c, where $S = 1$ and $F = N$, this is analogous to a crossbar switch (see section 6.3). Many optical interconnection systems limit fan-in to two (see appendix C), which implies that for shuffle networks to have a low blocking probability, many more than $\log_2 N$ stages are required.

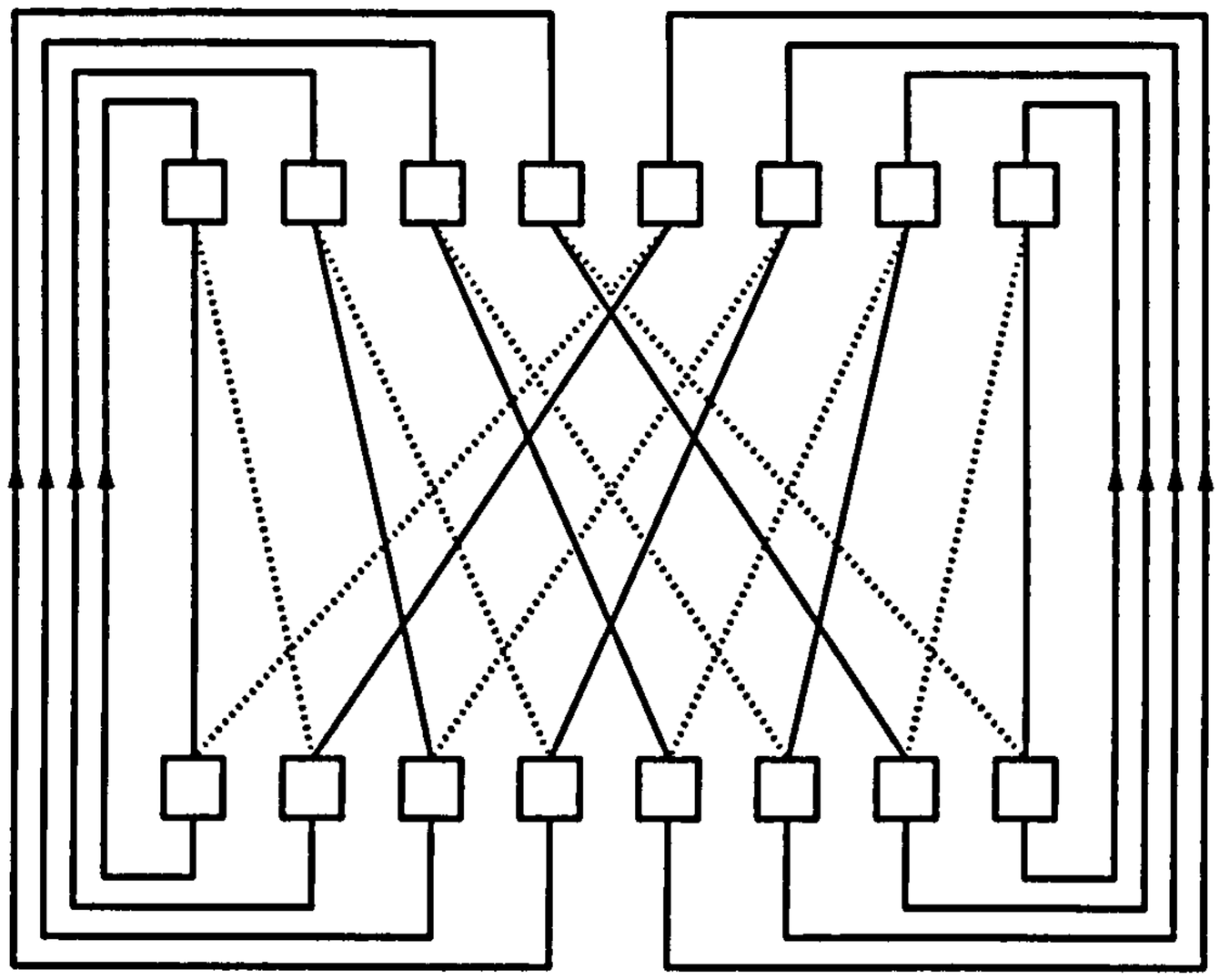


Figure 1.5: Schematic illustration of a perfect shuffling machine, an architecture for a general purpose parallel machine that takes advantage of the stage-invariance of the perfect shuffle network (see figure 1.3a). Data is clocked through the interconnection machine (via logic planes - not shown) the required number of times to implement an algorithm or network scheme.

1.3.4 The perfect shuffle

The large number of interconnection stages required to implement a logarithmic switching network for optical processing or communications can be prohibitive in terms of hardware cost, e.g. 256 channels requires 8 interconnection stages and 9 optical logic planes. The topology of the perfect shuffle is therefore particularly attractive because it is stage-invariant. This allows the network to be folded up on itself to form a perfect shuffling machine, illustrated in figure 1.5. The additional overheads incurred by this added flexibility are the reduced ability to pipeline data through the network and the extra clocking and control required to initiate communications across the net and extract data from it.

The perfect shuffle finds its true niche within parallel machine architecture [STI88]. It can be used to compute fast Fourier transforms (FFTs), sorts, merges, polynomial evaluation and matrix calculations [SED88, chapters 40 & 41], [STO71]. Murdocca [MUR88] has shown how the perfect shuffle (and topologically equivalent networks) can be used to implement programmable logic arrays (PLAs), symbolic substitution processors and hypercube networks (used in the implementation of an

optical version of Hillis's connection machine [HIL85]). The architectural versatility and scalable connectivity of the perfect shuffle have led to many optical implementations of the interconnect, e.g. [LOH86, BRE88, HAU89, ROB91a, WAN93, DES95, KOB95, PRI95].

1.4 System specification

In designing demonstration optical interconnection systems there are several important criteria to be considered. As optical logic devices have become smaller and systems have progressed towards faster optics (lower f/number), alignment tolerances have become critical and an area which is now of paramount importance is the optomechanical design and fabrication of the system [MCC93]. As far as holographic interconnects are concerned, this requires reproducibility and control over processing techniques, recording geometries and recording wavefronts. This is the one area where DCG volume HOEs have shortcomings. Although it is possible to fabricate highly accurate interconnects (as outlined in chapter 2), mass production/reproducibility is difficult to achieve and has led to the widespread preferential use of surface-relief diffractive optical elements [MIL93]. Newer volume media such as DuPont photopolymer [SMO89] should make this less of a problem.

1.4.1 System efficiency

In modern electronic or optical communication systems, where data transmission rates can be several Gbits/s, it is important to minimise the occurrence of errors without sacrificing speed/processing power. Consequently, a low bit error rate (BER) for the underlying communication system is important. BERs in the region of 10^{-6} are acceptable when error detection is coded into the data, e.g. cyclic redundancy check [DUN94]. This, however, adds to processing overheads required to detect errors and slows the system down. To make full use of the data bandwidth that is possible with optical communication systems, BERs $< 10^{-12}$ per channel are desirable.

Low BERs require a high signal-to-noise ratio (SNR) for the digital optical sys-

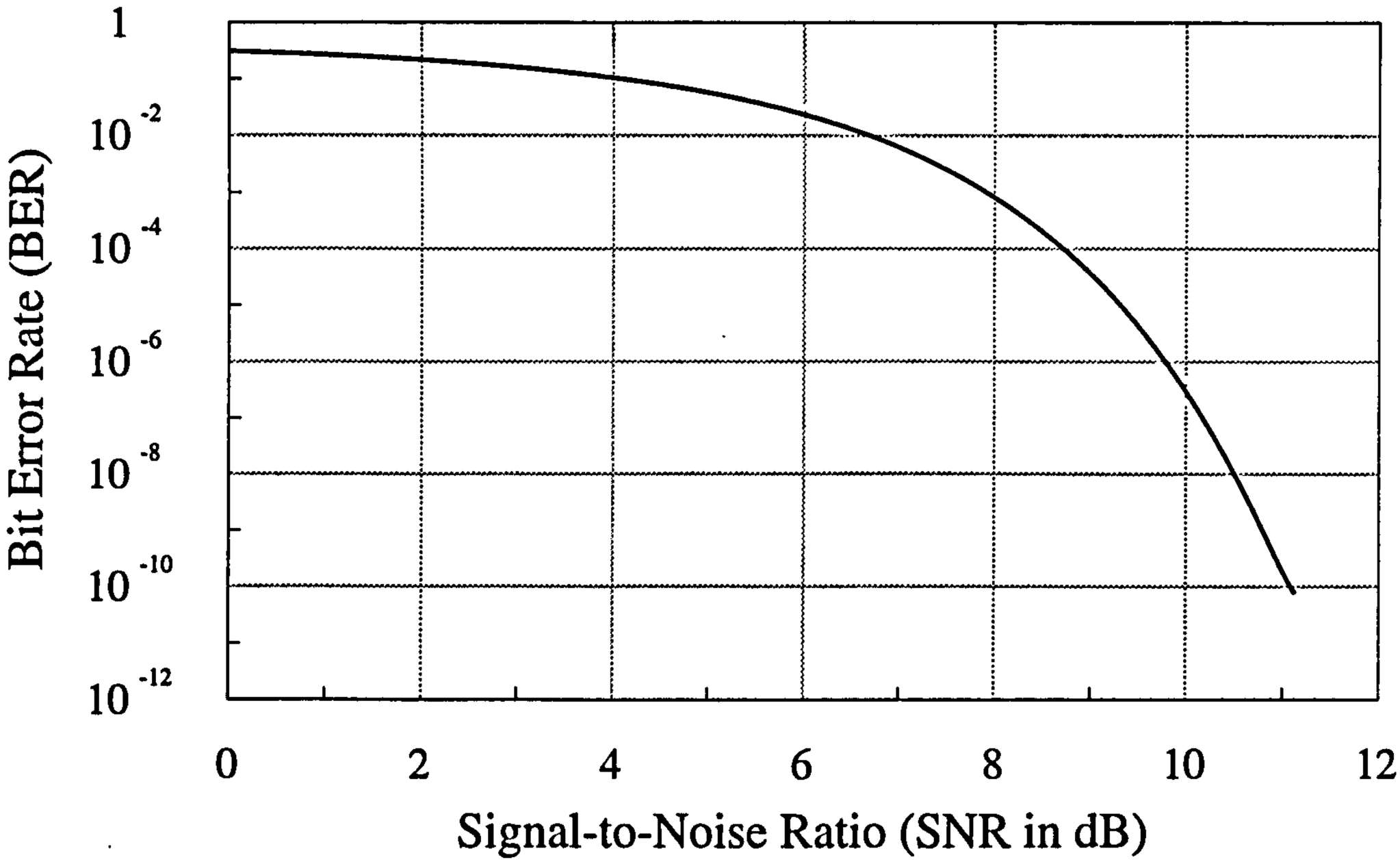


Figure 1.6: Bit error rate (BER) variation with signal-to-noise ratio (SNR), derived from statistical decision theory, after [DUN94], (the $y-$axis is logarithmic).

tem. In order to attain a large SNR both the efficiency and uniformity of power distribution across the field of the optical interconnection system should be as high as possible. For an optical power SNR defined as

$$\text{SNR} = 10 \log_{10} \frac{S}{N}, \tag{1.1}$$

where S and N are the signal and noise respectively, Dunlop & Smith [DUN94] show how this relates (through statistical decision theory) to the resultant BER, assuming a Gaussian white noise distribution:

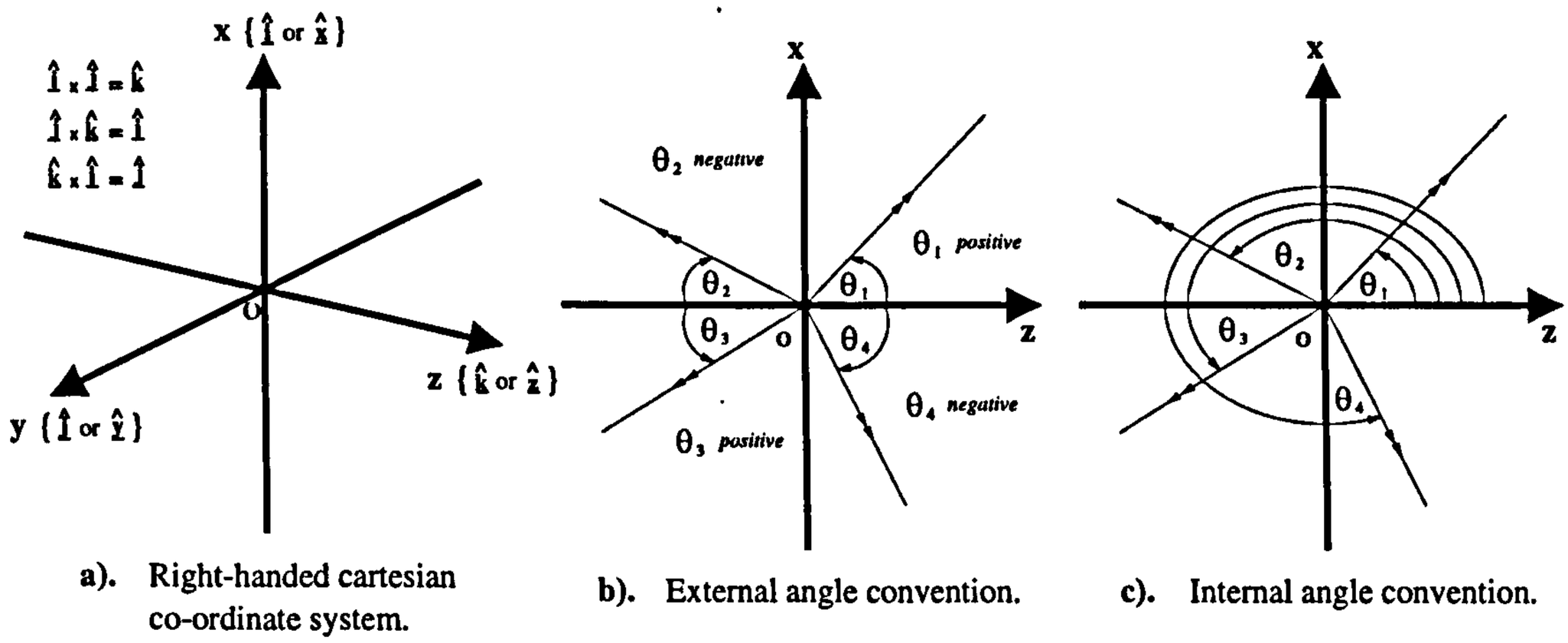
$$\text{BER} = \frac{1}{2}\left[1 - \text{erf}\left(\frac{S}{2\sqrt{2}N}\right)\right] \tag{1.2}$$

Figure 1.6 illustrates this dependence of BER upon SNR. It can be seen that for BERs below 10^{-6} small increases in SNR (above 10dB) can significantly reduce the BER. Consequently, the more efficient the interconnect the better.

1.4.2 Device considerations

All optical interconnects require some form of light source and optoelectronic device to perform their function. These devices dictate the operating wavelength of the interconnect, which for SEEDs and NLIFs is in the near infra-red (IR). Current semiconductor diode laser technology and vertical cavity surface emitting laser (VCSEL) arrays [JAH92] are capable of producing a maximum of only a few hundred mW of optical power at these wavelengths. This along with heat dissipation requirements makes it imperative that the optical logic elements consume a minimal amount of optical switching energy. Since switching energy is proportional to the active device area, this area needs to be very small (SEED windows are typically between 5 and 10μm). Consequently, optical interconnects need to produce efficient point-to-point connectivity in the near IR using highly focused beams. This requires fast optics (typically $\leq f/5$) working close to the diffraction limit.

Many optoelectronic devices, such as the symmetric-SEED (S-SEED) implement dual-rail (or differential logic); as opposed to the thresholding logic of devices like NLIFs. Thresholding logic devices are more intolerant to non-uniformity of power over an array of optical inputs [CRA91, page 34]. They therefore require a higher uniformity of efficiency and encircled energy between channels, which for holographic interconnects means a uniformity of diffraction efficiency.



a). Right-handed cartesian co-ordinate system.

b). External angle convention.

c). Internal angle convention.

Figure 1.7: The standard coordinate system and angle convention adopted throughout this thesis

1.5 Coordinate system and angle convention

The coordinate system and angle convention used throughout this thesis are shown in figure 1.7. A distinction is made between internal and external angle conventions with respect to the HOE:

- Angles, θ, external to the HOEs are (for convenience) defined relative to an axis perpendicular to the HOE surface ($z-$axis), shown in figure 1.7b (the HOE lies in the $x-y$ plane), $-90° \leq \theta \leq 90°$. The replay and recording angles of HOEs are usually defined between 0°and 90°; this angle convention defines which quadrant of the coordinate system beams are incident from.

- Angles, θ, internal to the HOE medium are (for mathematical modelling purposes) defined between $0° \leq \theta \leq 360°$ relative to the positive side of the $z-$axis. The external angle convention can lead to ambiguities when used in mathematical models.

1.6 Thesis layout

The thesis is constructed of six main chapters. Chapters 2 to 5 describe the methods, theories and techniques that have been used to produce a wide range of DCG HOEs during the course of this research:

- **Chapter 2** briefly describes volume HOEs and reviews why DCG is a good medium in which to record highly efficient volume holograms. The basic processing theory of refractive-index modulation formation is discussed and the pre-processing, recording, development and post-processing techniques used for interconnects throughout the thesis are described.

- **Chapter 3** describes a new volume hologram model used to predict and analyse transmission and reflection HOE characteristics throughout the thesis. The model is an extension of Moharam & Gaylords [MOH81b] rigorous coupled-wave theory (RCWT) and it is described in relation to the overall field of diffractive optical methods.

- **Chapter 4** addresses the problem of the restricted photosensitivity of DCG (discussed in chapter 2), presenting a new method of recording HOEs, both with and without focusing power, for replay at a wavelength beyond the sensitivity limit ($\sim$ 530nm). This is particularly important because most demonstration communications/processing systems operate in the near IR spectral region ($0.8 - 1.6\mu$m).

- **Chapter 5** extends the RCWT of chapter 3 to develop a new and generally applicable model for volume holograms multiplexed in the same emulsion (fan-out holograms). A discussion of fan-in and fan-out requirements and limitations of optical systems is presented together with an investigation concentrating largely upon reflection fan-out, (to parallel the transmission fan-out analysis of Redmond [RED89a]), which up until now has not been extensively investigated.

The methods and theories discussed in chapters 2 to 5 are then used to design, fabricate and analyse two distinct groups of optical interconnect:

- **Chapter 6** covers several new volume HOE implementations of spatially invariant interconnects (SIIs) together with their use in demonstration systems.

- **Chapter 7** presents a new and novel method of recording spatially variant interconnects (SVIs). Design considerations and experimental results are presented along with ways of incorporating them into demonstration circuits.

Finally:

- **Chapter 8** draws conclusions from the theoretical and experimental chapters regarding the current usefulness of DCG and volume HOEs applied to optical interconnection systems.

Chapter 2

Recording Holograms

A coherent recording can be made of the interference pattern between the scattered field of an object and a plane or spherical reference wave on a photographic plate. In 1948 Gabor [GAB48] demonstrated that if the latent image on the plate is developed and re-illuminated, by the reference wave, the recorded interference pattern would completely reconstruct the whole object wavefront by diffraction. This reconstruction process he called holography and the recorded interference pattern, a hologram; which is derived from the Greek meaning 'whole record'.

Holograms are devices that re-direct light into controlled directions, in contrast with a photograph which simply scatters light in all directions. This property of direction control forms the basis of the holographic interconnects which this thesis investigates. If a hologram is recorded between a plane reference wave and a spherical object wave, then replayed with the exactly the same plane wave, the spherical wave would be regenerated. This is the holographic equivalent of a lens and is an example of a Holographic Optical Element or HOE. HOEs can duplicate optical elements by recording the interference pattern between the desired input and output waves. A single HOE can form a holographic interconnect, but more usually a regular array of HOEs, perhaps several arrays coupled together, form the basis of the various interconnects discussed in this thesis.

HOEs fall into the two broad categories of transmission or reflection holograms each of which having very distinct characteristics and therefore different uses. Both

transmission and reflection holograms can be thick (volume) or thin and either amplitude or phase in nature. These various types of hologram have been discussed in many texts, e.g. those of Hariharan [HAR84] and Kogelnik [KOG69]. Transmission and reflection volume phase holograms are of interest in this work; they are discussed in more detail in chapter 3.

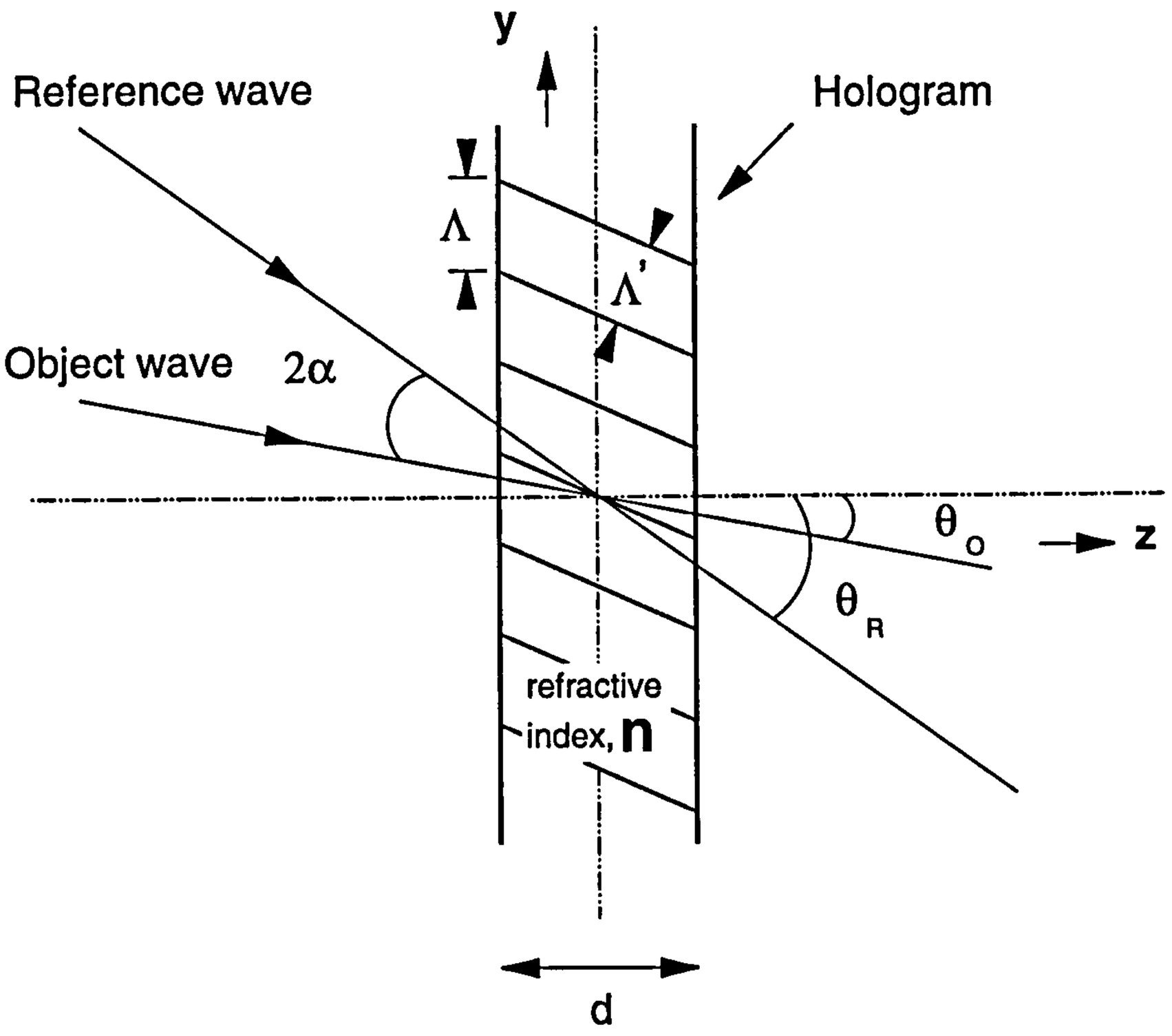


Figure 2.1: Arrangement for recording a transmission holographic grating structure.

2.1 Transmission hologram formation

A transmission hologram is formed by recording the interference between two wavefronts incident from the same side of a holographic plate. If, for example, two plane waves are incident on the holographic plate at reference and object angles of θ_R and θ_O respectively, as shown in figure 2.1, then a sinusoidal interference pattern, of period Λ', is recorded in the volume of the holographic medium. The maxima of this sinusoid are termed the fringe planes and they bisect the angle between the recording beams inside the medium. In a volume phase hologram this represents a

periodic refractive index modulation, the amplitude of which, Δn, determines the efficiency, η, of the HOE. The simple analysis of Kogelnik [KOG69] shows the following relationship between efficiency and the physical parameters, at Bragg incidence (see section 2.3), for the geometry illustrated in figure 2.1:

$$\eta = \sin^2\left(\frac{\pi \Delta n d}{\lambda \cos \alpha}\right) \tag{2.1}$$

This is a good working approximation when thinking about transmission hologram behaviour, however, it neglects many important physical effects, which require attention in order to make HOEs to a precise specification. These have been assessed by Redmond [RED89a] and will be reconsidered in chapter 3.

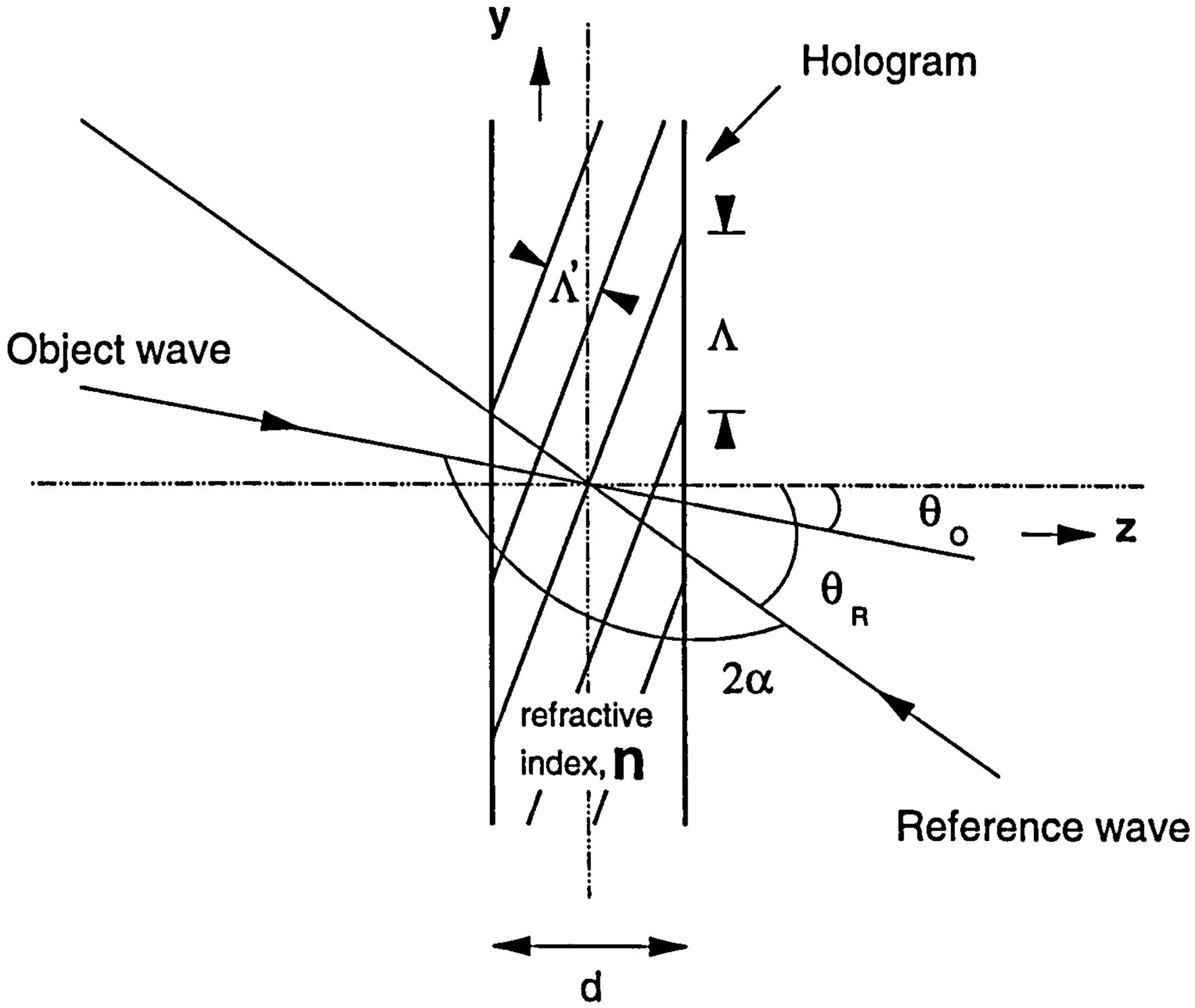


Figure 2.2: Arrangement for recording a reflection holographic grating structure.

2.2 Reflection hologram formation

In contrast to the transmission case, reflection holograms are formed by the recording of the interference pattern between two beams incident from opposite sides of the holographic plate, as shown in figure 2.2. Again, for a phase hologram, the fringe planes have a period Λ' corresponding to a refractive index modulation through the depth of the holographic medium. In this case, Kogelnik's simple analysis relates the diffraction efficiency to the physical parameters, at Bragg incidence, by:

$$\eta = \tanh^2\left(\frac{\pi \Delta n d}{\lambda \cos\alpha}\right) \tag{2.2}$$

Once again this is a reasonable working approximation (though not as good as in the transmission case). A more detailed assessment of the important physical effects in reflection holograms has been carried out by Newell [NEW87] and an accurate analysis of reflection HOEs is presented in chapter 3.

2.3 Replaying holograms

When a hologram is replayed with one of the wavefronts with which it was recorded, (at the wavelength at which it was recorded), it will reconstruct the other wavefront exactly (assuming no shrinkage or swelling of the holographic medium). This is called Bragg incidence and is embodied in the Bragg equation. For the transmission geometry of figure 2.1 the Bragg equation is:

$$\sin\alpha = \frac{\lambda}{2n\Lambda'}\,, \tag{2.3}$$

and for the reflection geometry of figure 2.2 it becomes:

$$\cos\alpha = \frac{\lambda}{2n\Lambda'} \tag{2.4}$$

If the HOE is replayed at a different wavelength or with a different wavefront the diffracted wavefront and diffraction efficiency will change. A detailed analysis of the

wavelength and wavefront performance of volume HOEs is presented in chapters 3 and 4.

2.4 The recording medium

There are many photosensitive materials which are suitable for recording volume phase holograms, (see [SMI77] & [HAR84, chapter 7]), although dichromated gelatin (DCG) and silver halide have to date proven the most useful. There have, however, been recent developments in photopolymer technology, most notably those of DuPont [SMO89] and Polaroid [ING89], which have led to new materials with undoubtedly an important role to play in future hologram development.

There are many considerations to be made when analysing materials which could possibly be used for recording holograms. However, when considering volume holograms to be used for practical holographic optical interconnects there are nine main factors to be weighed:

1. **Efficiency:** A volume hologram should only produce power in a specific set of diffracted orders (usually one if it is not a fan-out device, see chapter 5). In addition, it should be highly efficient in terms of redirecting optical power into these orders; ideally the zero order should contain no power, when used in the correct replay geometry. This requires that a large and controllable refractive index modulation should be achievable in the medium.

2. **Resolution:** An optically recorded hologram should be capable of faithfully replaying the wavefront it was designed for, in order to do this the medium should capture all the spatial frequencies in the recording wavefronts (many thousands of line pairs per millimetre, lp/mm), ideally this should be a molecular resolution.

3. **Uniformity:** The photosensitive layer should be of a uniform thickness, so that there is no wavefront distortion by the medium (or the substrate) or variation in volume holographic properties over a given area.

4. **Loss mechanisms:** Since efficiency should be a maximum, any absorption or scattering, during reconstruction, due to the medium is undesirable. Scatter is particularly unwanted because it can appear as noise on the recreated wavefront. Whereas absorption can lead to damage when reconstructing with high power beams.

5. **Sensitivity:** High sensitivity can be both beneficial and detrimental. It is desirable to have a material which is very sensitive in order to reduce exposure times (long exposures increase the risk of unstable interferometric recording). If, however, the medium is too sensitive it becomes difficult to handle without the risk of exposure. It is also important that the material, ideally, has a photoresponse from the ultra-violet to the near infra-red.

6. **Linearity:** A medium should show a linearity of response with exposure energy, which in photographic terms is called reciprocity. A medium which has good reciprocity should record a grating of equal amplitude for each sequential exposure of equal energy.

7. **Processing:** A medium should be capable of being processed in a relatively straightforward manner. In addition, the processing procedure should have no effect upon the physical parameters of the medium or the recorded holographic grating structure. The main undesirable effects are those of swelling and shrinkage of the medium. The processing procedure should be strictly controllable, allowing reproducible results.

8. **Damage threshold:** A hologram once processed to achieve peak efficiency should be able to operate in it's intended environment. In addition, if the absorption of a HOE is not zero, the temperature changes which can occur within it, during normal operation, can lead to damage. Changes in environment should be endured without loss of efficiency and they should not affect the replay parameters of the hologram.

9. **Lifetime:** A holographic medium should be able to retain the recorded HOE without deterioration due to aging effects.

Of the holographic materials available to date, the one which is capable of performing best in most of the above respects is DCG. It is not an ideal material but the shortcomings which it possesses can be circumvented.

2.5 Dichromated gelatin

Dichromated gelatin is a well documented holographic material, its use for making HOEs has been covered in great detail by Chang & Leonard [CHA79], Graube [GRA78], Redmond [RED89a] and Newell [NEW87] amongst others. This thesis will concern itself with the aspects of it's use which are important in order to understand how the interconnects described were fabricated.

Gelatin is derived from a naturally occurring protein, collagen, by a lengthy chemical process which breaks it down into polypeptide chains. It is this biological origin which makes gelatin tricky to work with because the exact composition, and consequent properties, vary greatly from source to source. The gel phase of gelatin is clear, rigid and crystalline and when hardened (so that it does not form a solution) can absorb up to three times it's own mass of water, swell and remain intact. This makes gelatin conducive to wet chemical processing procedures. Indeed it forms the host matrix for silver halide photographic emulsions.

2.5.1 The photochemical reaction

A gelatin film can be made photochemically sensitive by the addition of a chromate or a dichromate to the crystalline structure. The most widely used sensitising agent is ammonium dichromate (ADC) where the photochemical reaction takes place via the Cr^{6+} ion. On exposure to light the Cr^{6+} ions are reduced to Cr^{3+}, which then react with the gelatin molecules to initiate complicated crosslinks between neighbouring molecules. These molecular crosslinks are very strong and stable, and they effectively harden the gelatin, reducing it's ability to absorb water. The hardness variation within the gelatin affects the density and hence the local refractive index of the medium. The gelatin therefore retains a photo-induced image, in the form of a density variation, which can then be amplified into an efficient hologram (see

section 2.7.2). Since the photochemical reaction involves only a few chromate ions and adjacent gelatin molecules (with a typical molecule length of 80nm) the spatial resolution is almost molecular. It has been shown that DCG is capable of recording spatial periods smaller than 200nm [MEY77].

2.5.2 Modulation formation

The refractive index modulation available from the photochemical reaction described in section 2.5.1 is not sufficient to create an efficient volume hologram. The modulation must be amplified using a process similar to Shankoff's technique [SHA68] which involves the rapid dehydration of the washed gelatin layer by immersion in an isopropanol bath. When the gelatin layer dries a greatly amplified modulation image results, although the mechanism for this amplification is not completely understood. There are many theories to explain this amplification but the one which appears to be most accepted is that suggested by Chang [CHA80] and others [MCG80, CUL82]. This theory assumes that the rapid dehydration causes a differential swelling in the gelatin between areas of high and low water absorption. This swell between hard and soft areas is frozen by removing the water before the gelatin can relax leaving a greatly amplified refractive index modulation. This model is assumed throughout the processes used and developed in this thesis and appears to be able to account for all that is observed.

2.5.3 Gelatin hardness

On account of the processes outlined in section 2.5.2, above, to record highly efficient and reliable holograms in DCG it is necessary to be able to control the gelatin hardness. The hardness of a gelatin layer, as investigated by Samoilovich [SAM80], is measured according to how much water the gelatin layer can absorb. This is expressed as a swelling ratio, S, defined as:

$$S = (w_s - w_o)/w_o \tag{2.5}$$

Where w_s is the mass of the swollen gelatin layer and w_o is the mass of the dry gelatin layer. The ideal value of S for holographic gelatin is in the region 2 to 3. The bias hardness of gelatin layers can be controlled at various stages of the layer preparation to obtain the optimum level for holographic applications; as described in detail by Redmond [RED89a].

2.5.4 DCG as a holographic medium

Ammonium dichromate sensitised gelatin which is of the correct hardness can be used with wet processing techniques, (described in section 2.7.2), to meet the requirements outlined in section 2.4. The properties of DCG that make it such a good volume phase holographic medium are:

1. Diffraction efficiencies greater than 95% are achievable for transmission holograms and in excess of 99% for their reflection counterparts. This corresponds to a maximum refractive index modulation of $\Delta n \sim 0.08$. The modulation is controllable by varying the processing techniques.

2. A spatial resolution close to the molecular scale is available with greater than 10^3 lp/mm being possible, indeed it is capable of recording spatial frequencies up to at least 5000 lp/mm [CHA79].

3. Gelatin layers can be cast, spun or obtained commercially up to 120 μm thick with an optical flatness better than $\lambda/8$ before and after processing.

4. Processed DCG holograms have scatter and absorption losses which are wavelength dependent. Typically an absorption loss of ~4% is observed at both 488 and 514.5nm and scatter is marginally higher at 488 than at 514.5nm. The level of scatter is greatly dependent upon the processing and bias hardness of the gelatin but in most circumstances is negligible.

5. The spectral sensitivity of DCG spans from the ultra-violet (UV) to about 580nm (falling off rapidly beyond 530nm), the peak being close to 488nm. This makes it easy to handle in the dark, using a red safelight but makes it

impossible to record HOEs directly for the red and near infra-red (IR). The two main recording wavelengths are 514.5 and 488nm because they are close to the peak sensitivity of DCG and they correspond to the two strongest operating frequencies of an argon ion laser. The exposure sensitivity of DCG varies with recording wavelength, processing procedure and the source of gelatin, but typically is in the range 10^{-4} to $10^{-5}m^2/J$ at 488nm and goes down to a fifth of this at 514.5nm. This is a factor of about 10^{-3} to 10^{-4} less sensitive than silver halide but, provided an argon ion laser of a few watts power is available, it does not lead to inordinately long exposure times.

6. Over the period of nine years that DCG HOE fabrication has been pursued in the physics department at Heriot-Watt there appears to be no evidence of reciprocity failure.

7. DCG has to be wet processed, which will be explained in section 2.7.2. This can cause differential bulk swelling across an HOE area, leading to volume holographic and geometrical wavefront non-uniformities, (by way of diffraction from the resulting surface relief structures), in replay. However, by carefully controlling all the environmental and processing parameters, reproducible and uniform HOEs can be produced. What is more, post-process tuning of volume parameters is possible, and reprocessing can recover humidity damaged holograms. However, several authors have reported evidence of such a failure [CHA79, PAS92].

8. Fluctuations in temperature and humidity can destroy DCG holograms [NAI90]. However, methods do exist to increase the damage thresholds to military specifications and sealing the gelatin layers before they leave the controlled production environment makes the HOEs stable [ROB91b, RED89a].

9. Holograms recorded in DCG, after being sealed from the environment, appear to show no aging or other deterioration effects.

The limited spectral sensitivity of DCG can be overcome in many applications of HOEs for the near IR by using optimised recording geometries in the visible; such techniques will be described in chapter 4.

2.6 Gelatin layers

DCG layers have been deposited in-house and a variety of readily available layers have been treated to make them suitable for holographic recording [RED89a]. All of the HOEs made in this thesis were recorded on gelatin layers made available to us by Pilkington plc, or those derived from Kodak 649F photographic emulsions.

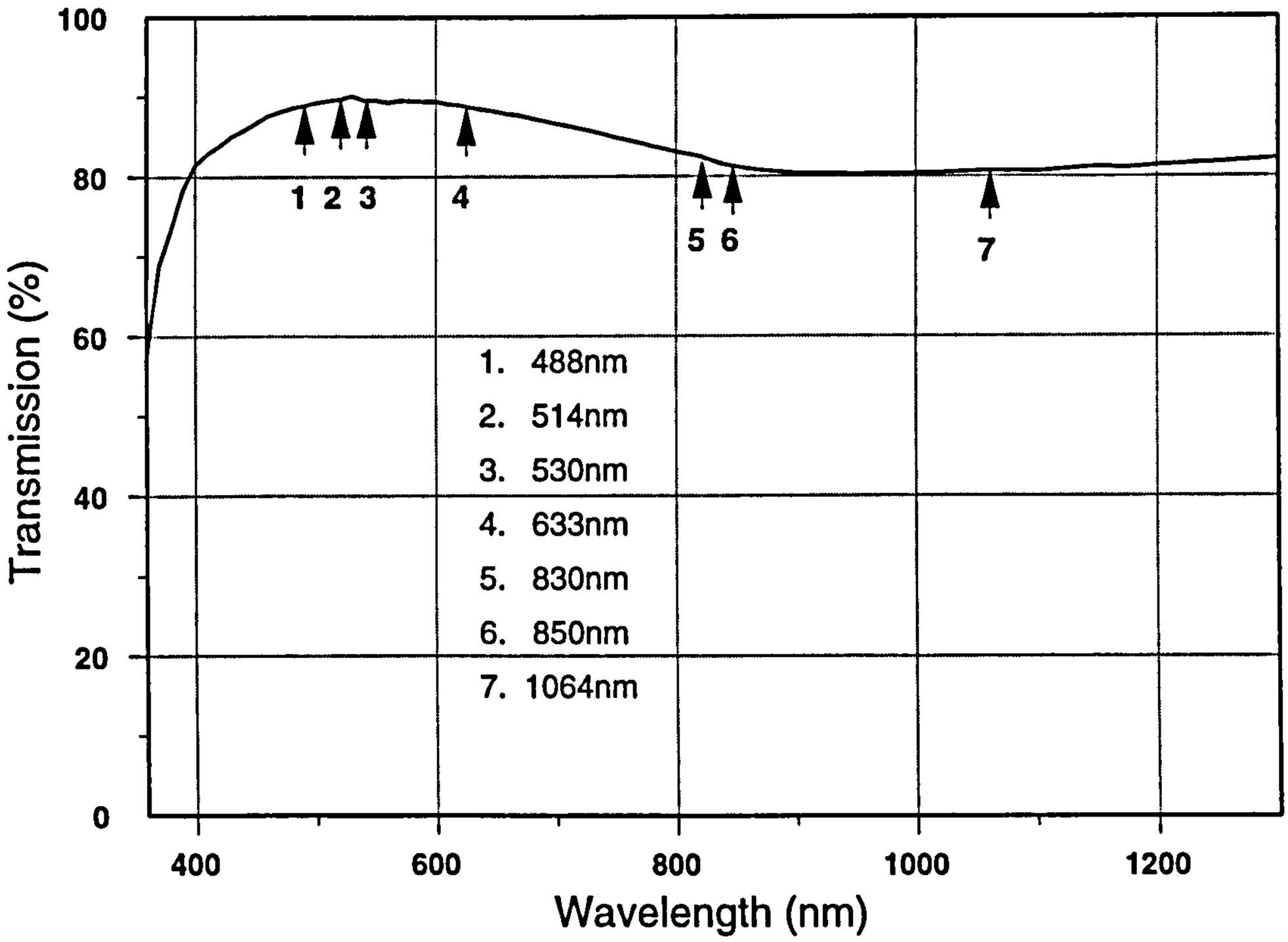


Figure 2.3: Transmission spectrum of a 40μm Pilkington layer and substrate, (arrows indicate typical HOE operating wavelengths in the visible & near IR).

2.6.1 Pilkington layers

A range of holographic quality gelatin layers have been provided to the holography group by Pilkington plc, with a choice of 5, 10, 40 and 60μm thicknesses, all on 2mm

thick glass substrates. These have a typical swell factor, S, of 2 and were delivered unsensitised but pre-hardened. A typical transmission spectrum of a 40μm thickness layer and the substrate is shown in figure 2.3, included in this is a Fresnel reflection loss of approximately 8% from the two surfaces of the gelatin and substrate; this spectrum is modified slightly by processing. The Pilkington plates have only been used in this thesis material where very large refractive index modulations have been required, as in the case of reflection holograms such as those in chapter 6.

2.6.2 Kodak layers

The large majority of experiments carried out in this thesis have been performed on Kodak 649F spectroscopic plates. These are a panchromatic photographic emulsion consisting of a 15μm thick gelatin matrix sensitised with silver halide grains on a 1mm thick glass substrate. These layers have to be preprocessed to remove the silver halide before they are suitable for DCG holographic purposes, the process used is based on that of Chang [CHA71] and is summarised in table 2.1.

Kodak 649F preprocessing		
Lighting: darkness.		
1. 25% Agfa g334 fixer (part A)	10 minutes	20°C
Lighting: room light.		
2. Water wash, increasing the temperature from 20 to 32°C at 1.5°/min.		
3. Water wash	15 minutes	32°C
4. 25% Agfa g334 fixer plus 5% hardener (part B)	10 minutes	20°C
5. Water wash	15 minutes	20°C
6. Dry in circulating air	overnight	20°C

Table 2.1: Process for removing silver halide from Kodak 649F gelatin layers.

The hardening step (step 4) can be carried out at various concentrations to yield swell factors in the range 2 to 2.5. The gelatin layers on Kodak plates are invariably too soft as they are supplied and can be hardened to the level required for the particular holographic application. This is not true for gelatin layers derived from other manufacturers of photographic emulsions, such as Agfa whose plates have typical S values in the region of 0.25, these layers need softening which considerably

complicates the preprocessing stage [OLI84].

Kodak 649F plates that have been preprocessed (following the steps in table 2.1), but have been found to be too hard for a specific application have been successfully softened by soaking in water at a temperature above that at which they were preprocessed. Figure 2.4 shows the results of such a soaking process on a batch of similarly preprocessed plates. The plates were soaked in water for 10 minutes at 20°C, then this water temperature was increased at 1.5°C/min to reach the required soaking temperature in each case shown. Each plate was then soaked at this temperature for a further 10 minutes. It can be seen that increasing the soaking temperature partially melts the gelatin and consequently reduces the layer thickness. Soaking at temperatures above 50°C results in the gelatin having an unacceptable scattering level, (well in excess of 10%).

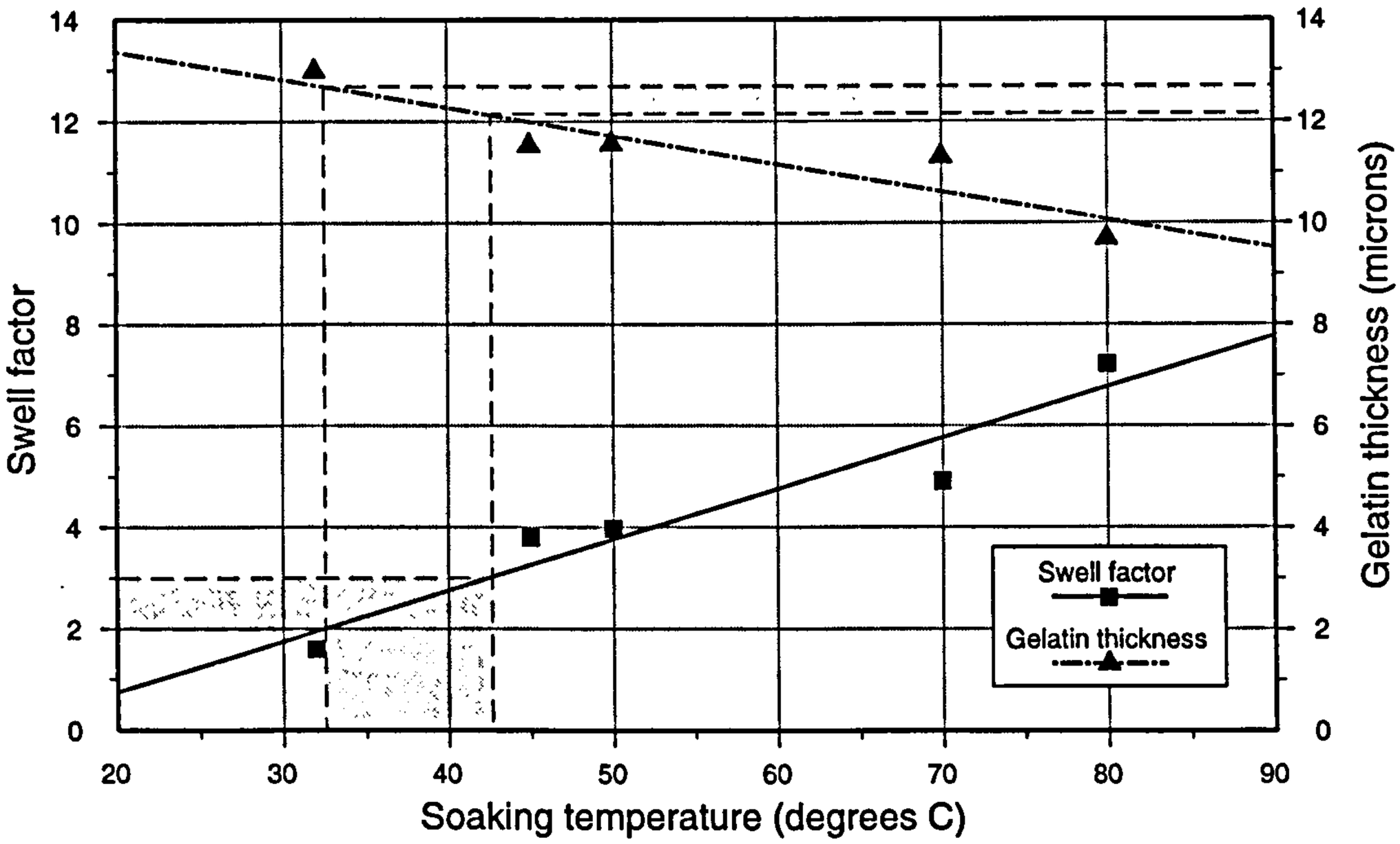


Figure 2.4: The effect on both gelatin thickness and swell factor of soaking an, already, preprocessed Kodak 649F plate at different temperatures. The empirical lines of best fit are shown for interpolation purposes. Shaded areas show the typical soaking temperature range (and resultant gelatin thickness) over which a usable swell factor can be achieved.

Swell factors in the region of 2 to 3 are highlighted in figure 2.4. The accompanying thickness change in this area is less than 1μm, which leaves a suitable gelatin

thickness for volume phase holography.

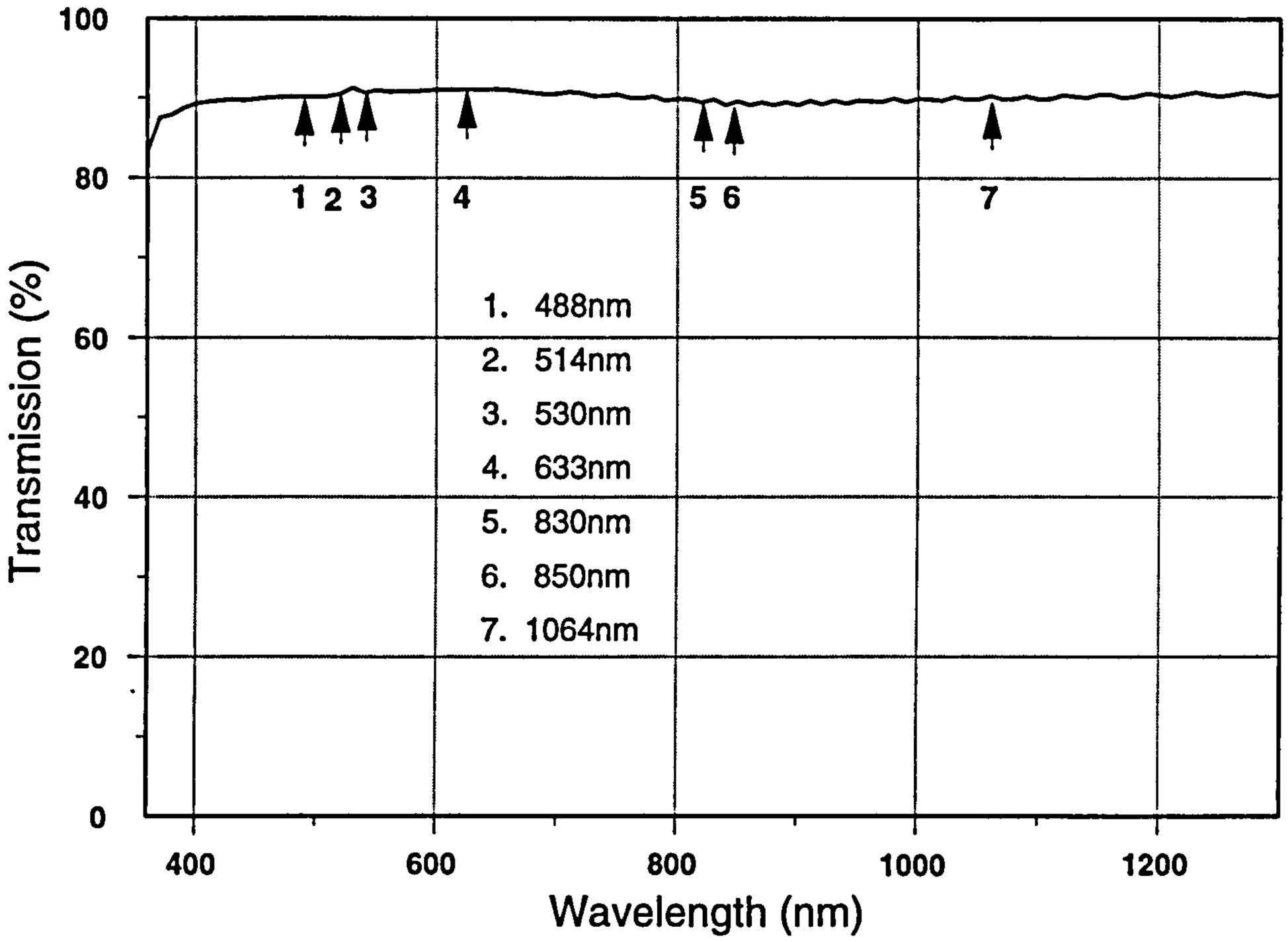


Figure 2.5: Transmission spectrum of a 15μm Kodak layer and substrate, (arrows indicate typical HCE operating wavelengths in the visible & near IR).

A typical transmission spectrum of a 15μm thickness preprocessed 649F layer and the substrate is shown in figure 2.5 (including Fresnel reflection losses); again this spectrum is modified slightly by processing. The transmission spectra, of both Pilkington and Kodak plates, show that the absorption, (after accounting for reflection losses), across the visible and near IR spectrum is low enough to make these layers useful for the holographic interconnect applications considered in this work.

2.7 Using DCG layers

The DCG processing steps are very sensitive to moisture content, especially when the plates are in the sensitised state and during processing. Consequently, the humidity and temperature must be strictly controlled throughout all the processes outlined in this section. The environmental parameters that give reproducible high quality

DCG HOEs are a temperature of 20°C ±2°C and a relative humidity (RH) of 40% ±5%. In this work this is achieved by having an air conditioned laboratory, however, it is possible to attain similar results using a laminar-flow table [LIN92]

2.7.1 Sensitising

Sensitised plates should be used within 24 hours when stored at 20°C and 40% RH, otherwise their sensitivity decreases rapidly. DCG undergoes a dark reaction which limits the shelf life, [NEW87, CUL82], therefore for consistent, reproducible results the plates should be sensitised and used within similar time spans.

The 15μm gelatin plates are usually sensitised (under dim red safelight conditions) in a 5% ammonium dichromate/water solution for 3 minutes at 20°C, and the 40μm plates in 2.5% ammonium dichromate/water for 5 minutes at 20°C. After sensitisation the plates are left to dry (overnight), in darkness, in circulating air at 20°C and 40% RH. The ammonium dichromate absorbs light during exposure and this absorption is directly proportional to the sensitiser concentration. The level of absorption, therefore, alters the effective thickness of the hologram during exposure and must be controlled to use the full thickness of a volume HOE to the best advantage. If concentration levels are too high the effective thickness of the medium in which a hologram is recorded will be reduced because most of the absorption will take place in the first few microns of the medium. This is where the most strongly modulated grating will be recorded, the effect of this is that the grating is not recorded in the available volume and consequently may produce more than one diffracted order and the achievable index modulation may be reduced [CHA79]. The sensitiser concentrations that seem to work the best with the 15μm and 40μm layers are 5% and 2.5% respectively.

2.7.2 Processing

Throughout this thesis a standard processing procedure has been adopted for use with the Kodak 649F plates, which are predominantly utilised. This processing procedure is shown in table 2.2 and is based upon that of Chang and Leonard [CHA79].

Standard processing procedure		
Lighting: dim red safelight.		
1. 0.5% $(NH_4)_2Cr_2O_7$ (Ammonium dichromate)	5 minutes	20°C
2. 0.5% $Na_2S_2O_7$ (Sodium metabisulphite)	5 minutes	20°C
3. Water wash	10 minutes	20°C
Lighting: room light.		
4. 50% Isopropanol/water	3 minutes	20°C
5. 75% Isopropanol/water	3 minutes	20°C
6. Isopropanol	3 minutes	20°C
7. Dry in circulating air	$\geq$1 hour	20°C

Table 2.2: The standard post-exposure DCG processing procedure for Kodak 649F plates.

The main problem with using gelatin layers derived from photographic emulsion is that the physical properties can vary between batches. In the extremes of this variation the modulation available in the gelatin from standard processing can be too low for certain applications. For example, when wavelength shifting to the near IR, (see chapter 4), the modulation required for maximum efficiency is considerably larger than for the same efficiency at the recording wavelength. Consequently a variation in processing to achieve the higher modulation, ("wash processing"), is sometimes required, as shown in table 2.3.

Wash processing procedure		
Lighting: dim red safelight.		
1. Water wash	$\geq$50 minutes	20°C
Lighting: room light.		
2. 50% Isopropanol/water	3 minutes	20°C
3. 75% Isopropanol/water	3 minutes	20°C
4. Isopropanol	3 minutes	20°C
5. Dry in circulating air	$\geq$1 hour	20°C

Table 2.3: Post-exposure wash processing procedure for higher index modulation.

The holograms that are wash processed have a lower absorption level (less than 1% at 514.5nm) but a higher scatter, this is the price paid for the modulation increase. Furthermore, if wash processing at 20°C does not yield a sufficiently large

refractive index modulation, the processing steps 1 to 4 above can be carried out at a higher temperature up to a limit of about 32°C. The higher the processing temperature, the higher the modulation and the greater the level of scatter; taking the processing temperature beyond 32°C may wash away holographic quality gelatin. This temperature limit depends upon the bias hardness of the gelatin and the processing used, but 32°C is a good guideline. If the processing temperature is taken too high for a particular gelatin layer the stresses on individual gelatin molecules can become too great and cracks may occur in the hologram on a microscopic scale as the fringe planes shear from one another. This turns the gelatin into a milky opaque layer with very high scatter levels and poor holographic qualities.

The Pilkington plates, being thicker than their Kodak counterparts, require longer processing times than given in the above two processing schemes. Typically 40μm layers require on average three times the processing times for each stage of the above processes.

The transmission spectrum of gelatin layers and substrates is altered a little by processing and by the type of processing procedure used, but not significantly from that shown in figures 2.5 and 2.3. For exact details see [RED89a].

2.7.3 Reprocessing

The most usual form of hologram degradation when left in an unsealed state is humidity damage. Holograms which have absorbed a lot of water have reduced index modulation but can be reprocessed to regain the former diffraction efficiency. The cross-linking formed in the photochemical reaction is not destroyed by the reabsorption of water or by normal processing. By repeating the washing and dehydration steps of a processing procedure it is possible to regain the index modulation, as described in section 2.5.2. Some degree of efficiency degradation occurs upon reprocessing unless the hologram is reprocessed within a few hours, and the level of scatter increases with every subsequent reprocessing stage [CUL82]. Increasing the temperature of the reprocessing stages, over that at which the original processing was carried out, increases the sensitivity of the hologram and enables the efficiency

degradation suffered to be reversed, with the preceding stipulation of increased scatter [CHA76].

Sealing the gelatin layers, with a glass plate and optical adhesive, fixes their replay parameters, see section 2.8.1. Often in an experimental environment this is undesirable because it may be required to tune several HOEs together to make a system of coupled holograms. These holograms may not have been fabricated at the same time but over a period of experimentation. Consequently, between recording and the final coupling experiment holograms will deteriorate somewhat. Leaving HOEs in an unsealed state and letting them deteriorate is an acceptable method of storing holograms until they are to be used. When they are required, all the holograms to be used in the interconnection system can be reprocessed and tuned, (as described in the next section), so that they couple to one another in the correct manner. However, it is preferable to store unsealed HOEs in a desiccator to avoid the extra work and possible performance degradation.

2.8 Post-process tuning and protection

Holograms can be damaged by environmental and optically induced thermal effects. In addition, after processing the gelatin layer is seldom the same thickness as it was before sensitisation. These factors can be controlled in DCG by using various baking techniques to tune and protect the HOEs. Baking a hologram has three main effects:

1. Water is driven off from the molecular complexes within the gelatin causing an increase in the melting point,

2. The gelatin layer thickness shrinks, because water is driven off; and

3. The gelatin is partially melted.

When the gelatin layer melts the refractive index modulation is reduced due to the thermal destruction of cross-links. The effect of driving off water tends to counteract this modulation reduction. The water a gelatin layer contains when it is in

equilibrium with the atmosphere cannot be removed by the processing procedures described in section 2.7.2. Driving off this water by baking is effectively a continuation of the dehydration step of processing and as such enhances the refractive index modulation (in accordance with the model summarised in section 2.5.2). The melting point of the gelatin layer is also increased by this additional dehydration since it is dependent upon the water content of the gelatin, being lower when more water is present.

It is, therefore, possible to protect DCG HOEs against thermal damage [ROB91b] and the diffraction efficiency can be tuned to a desired value [CHA79, RED89a] by baking.

The efficiency of a hologram can be optimised by initially over-exposing, typically by 50%, and then baking by gradually increasing the temperature from 70°C to 100°C, thus reducing the modulation to that required for peak efficiency. The damage threshold of the hologram is also increased, by this baking, so that it can stand temperatures in excess of 110°C [CHA79] and optical powers in excess of $2kW/cm^2$ in an unexpanded 514.5nm laser beam. This optical damage threshold can be in excess of 70 times the unbaked value [ROB91b]. There are a very large number of variables in the complete DCG process and it is still found that holograms treated identically at every stage of the process do not always yield the same results. Consequently the power handling and tunability can change between gelatin holograms.

The power handling of an HOE can be further increased by thermally bonding a hologram to a substrate of greater thermal conductivity, K_s. For example, Sapphire has a K_s of $27.21 Wm^{-1}K^{-1}$ compared to float glass with a K_s of $0.85 Wm^{-1}K^{-1}$, and this can increase an HOE power handling by up to 20 times [ROB91b]. Reflection holograms can be thermally bonded to materials such as copper to significantly boost the power handling beyond present limits.

A further effect of baking is shrinkage. It is very useful in altering the replay wavelength and/or angle of an HOE. A thickness change can be successfully produced without a change in efficiency by baking just under the melting point of the gelatin. Typically baking at 65°C for a period of 10 minutes to one hour will produce

a replay angle change between 1° and 2° (in a transmission HOE with 0° and 30° beam angles). Replay angle changes up to 4° can be achieved by baking a hologram at 120°C to 150°C for 10 to 30 minutes. This method rapidly drives off most of the water in the gelatin and reduces the layer thickness somewhat, consequently it will alter the efficiency as well as the replay angle.

Throughout the work in this thesis, wavelength and angular replay parameters have been successfully tuned using the baking processes described. The thermal damage threshold of the HOEs has been increased to beyond 100°C; for a detailed mathematical analysis of these baking processes and the other effects of heating holograms see Robertson et al. [ROB91b].

2.8.1 Sealing HOEs

Once a HOE has been processed and tuned to achieve the desired geometrical and volume replay parameters it can be sealed against the environment. Robertson et al. [ROB91b] show that the reabsorption of water into an HOE is initially very rapid, typically reabsorbing half the equilibrium value in only 10 minutes. Sealing an HOE must therefore be carried out before it starts to reabsorb water from the surroundings. Typically a hologram is sealed on a hot plate at an ambient temperature below the gelatin melting point but sufficiently high to prevent water from being reabsorbed, e.g. 40°C.

When sealing a hologram the gelatin layer is sandwiched between the substrate and a 1mm thick anti-reflection (AR) coated glass plate, (coated for the replay wavelength). The optical cement used is a UV curable adhesive which is not absorbed by the gelatin layer, made by Norland Products Inc. Once the HOE has been sealed with a cover slip the gelatin is cut away from the edges of the substrate and UV curable adhesive is applied along the join with the cover plate to prevent lateral diffusion of water into the hologram. The hologram can then be removed from the protective environment of the laboratory. Note that a sealed hologram will be more susceptible to thermal damage for a given water content, because the water cannot escape. This means that the melting temperature of the gelatin cannot rise as it

would if the water were able to be driven out.

2.8.2 Recording source

All HOEs described in this thesis were recording using a Coherent Innova 90/6 argon ion laser with a temperature stabilised Fabry-Perot etalon for single frequency operation at TEM_{00}. Without the etalon the laser has an effective bandwidth of 1GHz which implies a coherence length of 3cm, however, with the etalon this bandwidth decreases to 5MHz giving a coherence length in excess of 3m. This coherence length is large enough such that exact matching of the optical path lengths between the split arms of a typical recording scheme, (see figure 2.6), is not a critical factor in achieving optimum fringe contrast at the recording plane. This laser delivers most power (in excess of 1W) at 488 and 514.5nm which are ideal wavelengths for recording holograms in DCG.

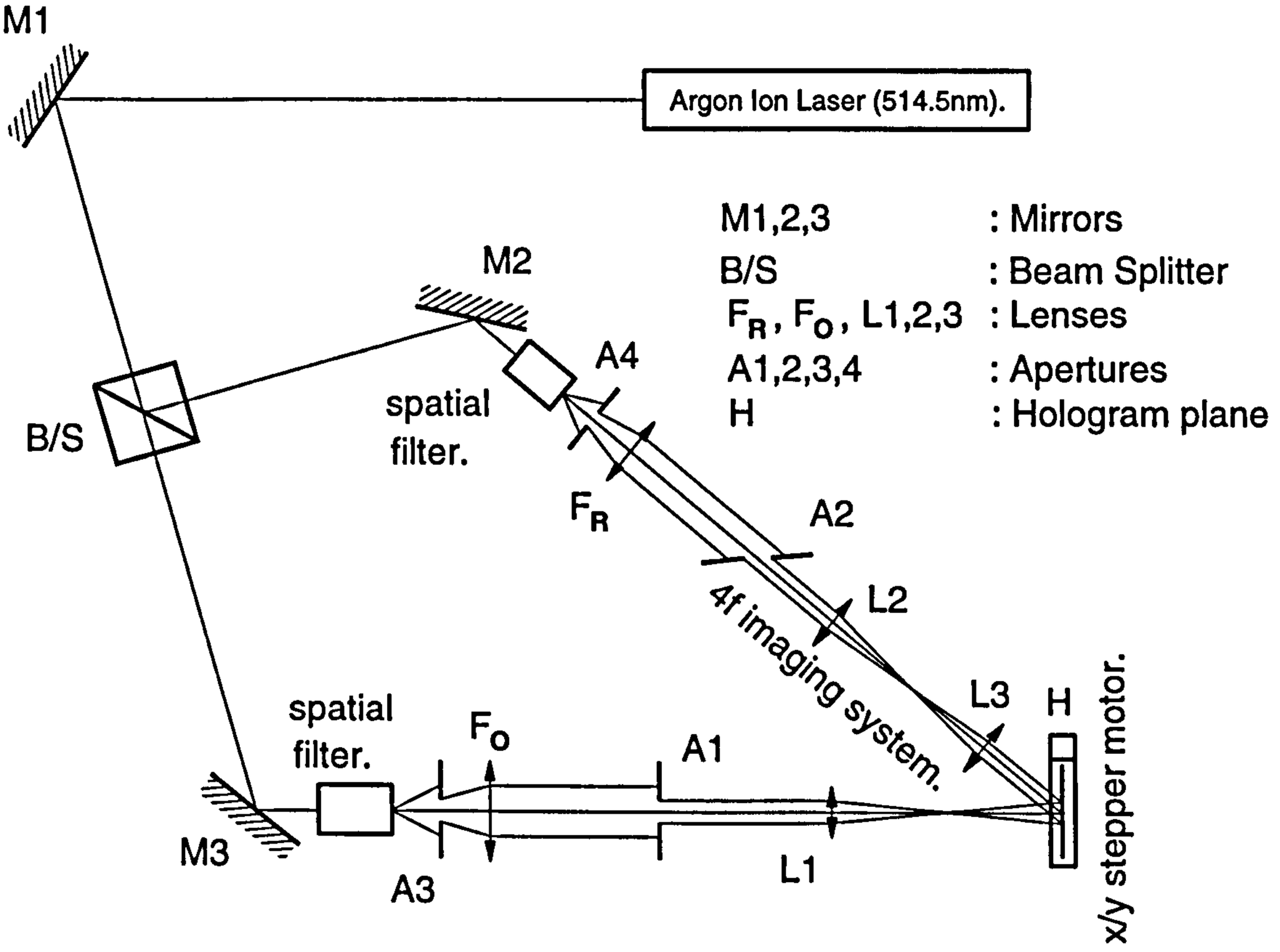


Figure 2.6: A typical transmission HOE interferometric recording geometry.

2.9 HOE recording procedure

A typical transmission HOE recording geometry is shown in figure 2.6. This is a split beam interferometric recording technique and as such requires interferometric stability during an exposure. To ensure vibration isolation, all HOEs are recorded on a floating optical table and exposure times are kept to a minimum. The effects of air movement and thermal settling of optical components are also minimised. The laser is isolated from the table by foam pads to avoid vibration due to the cooling water. The air conditioning unit, in the lab, is always switched off at least ten minutes before exposure to allow the air to settle and to permit the optics and recording plate to reach thermal equilibrium with the surroundings. Where necessary, curtains are drawn around the table during exposure to minimise air currents across the recording system. For long exposures, typically in excess of 5 seconds, the air conditioning is switched off for an hour or more before recording. Whenever possible irradiance levels are kept to below $500W/m^2$ to avoid thermal expansion of the gelatin during recording. All optics are AR coated wherever possible to avoid unwanted interference effects which will introduce noise modulation onto the amplitude of the desired holographic fringes. Dielectric mirrors are always used, to eliminate thermal expansion during exposure (which can be as high as $\lambda/10$ for aluminium mirrors in a typical unexpanded recording beam).

When recording reflection holograms the interferometric stability of the optical system is even more critical than in the transmission case. This is because the fringes are recorded almost parallel to the surface of the gelatin (as shown in figure 2.2) through the depth of the medium. Consequently any movement of optical components during recording can move the physical location of a fringe through more than the period of the recorded grating, wiping out all of the modulation.

2.9.1 Recording clean holograms

The Innova 90/6 produces a good quality Gaussian beam at modest intra-cavity aperture sizes, allowing at least 1W of vertically polarised (perpendicular to the optical bench), TEM_{00} light at 514.5nm. This does, however, have some high fre-

quency noise upon it, so to record the cleanest HOEs possible the split beams are spatially filtered then collimated. Only a small central portion of this collimated beam is used in recording to ensure a good approximation to a uniform amplitude wavefront. This in turn helps to ensure a consistent modulation across the area of a hologram, implying a uniform efficiency.

A typical recording arrangement uses imaging systems to generate the required recording wavefronts and to simultaneously sharply image apertures onto the holographic plate. This enables a very high packing density of HOEs to be achieved. When recording arrays of holograms onto a single plate, apertures down to 100μm diameter have been successfully recorded. High quality achromatic doublets or a matched pair of plano-convex lenses (for 4f imaging) are used in the imaging systems wherever possible. This helps to minimise wavefront aberration during recording and subsequent HOE replay.

To record arrays of holograms the holographic plates are mounted on a $x-y$ stepper motor stage which together with an electromagnetic shutter are under the control of a personal computer (PC). This allows automatic recording of test plates and arrays of HOEs with positional accuracies better than $\pm 10\mu$m. In addition, a kinematic plate holder has been designed to hold a variety of sizes of holographic plate (Pilkingtons and Kodak). This holder is magnetic and designed to place any holographic plate in the same plane in the recording system. It is important not to mechanically stress the holographic plate during recording for two reasons. The first is that when replayed the stresses cannot be exactly replicated and so the HOE will not reproduce the desired wavefront. The second is that if the plate is stressed it will not lie in the same plane of the recording system as it is stepped around. Since the recording system images very small apertures, typically 200μm, onto the holographic plate, it then becomes impossible to register the apertures of the two split beams across the area of the plate. This kinematic plate holder resolves both of these issues.

2.9.2 Index matching during recording

It is important to avoid reflections from the surfaces of a holographic plate (which can cause spurious gratings) during recording. To do this in the case of a typical transmission hologram such as that in figure 2.6, requires a neutral density (ND) filter, of at least ND 1, to be index matched to the back of the plate. This should be done with a suitable index matching liquid. Unfortunately most liquids of ideal refractive index (1.48 to 1.52), and of the correct viscosity, are toxic. Xylene is the normally cited index match, but this was considered too dangerous. For the majority of this work Di-n-butyl-phthalate has been used, again this is not a particularly safe chemical. Other alternatives include liquid paraffin, glycerol and cedarwood oil, all of which have been tried but all have their own specific problems. In future work, however, liquid paraffin, with a refractive index of 1.483, would seem to be the safest.

Some recording arrangements in this thesis require optical components to be index matched to the gelatin surface of the holographic plate during recording. Before the hologram is processed the index matching fluid should be removed, this is achieved by soaking the exposed plate in reagent grade methanol for 2 minutes (under safelight conditions) at 20°C.

2.9.3 Beam ratio

The split beam recording technique enables the control of the irradiance level in each arm of the system, this is most easily done using Holographic Attenuators (HAtts), first proposed by Redmond and Taghizadeh [RED88]. These are simple planar volume phase gratings, (recorded in DCG, and discussed in chapter 4), which provide a laser beam attenuation up to 18dB with a minimal beam deviation. HAtts are preferred over ND filters because they diffract not absorb and so are not susceptible to expansion and hence introduce minimal wavefront distortion during recording. Ideally the beam ratio, or k-ratio, is 1 to maximise fringe visibility [HAR84] and make full use of the available index modulation [CHA79, GEO87]. Increasing the k-ratio reduces the maximum achievable modulation, because the increased bias level of exposure due to unequal recording beams lowers the dynamic range of the refrac-

tive index modulation [BLA90]. There are, however, situations where a different beam ratio is required [BAR82], such as copying schemes for computer generated holograms (CGHs).

2.10 Summary

Efficient volume phase transmission and reflection HOEs are the basic elements of holographic optical interconnects. The best medium, to date, available for recording such HOEs is DCG. These HOEs can be recorded by a direct optical interferometric technique in DCG using the processes described in this chapter.

DCG can exhibit problems of poor reproducibility and sensitivity to environmental conditions (before sealing). The use of an environmentally controlled laboratory in which holograms are kept throughout all the stages of exposure and processing largely eliminates these problems. Under these conditions DCG becomes a manageable holographic medium which can be controlled in it's response by the processing and exposure techniques described.

DCG holograms that suffer humidity damage in an unsealed state can be reprocessed to regain former efficiency. Many of the final properties of a hologram such as efficiency, thickness and bandwidth can be controlled by baking in a conventional oven or on a hot plate. This baking also increases an HOEs damage threshold. Some experimentation has been carried out with microwave drying which can further increase an HOEs resistance to damage. A full study of this drying technique has been carried out by Rebordao and Andrade [REB89].

Once an HOE has been processed and tuned to obtain the desired properties it can be maintained in this condition by sealing. Sealing HOEs makes them stable outside of the laboratory and therefore practical.

Using the processes described, many accurately controlled HOEs have been recorded and assembled into optical interconnection systems for operation in the visible and near IR.

Chapter 3

Rigorous Coupled-Wave Theory

To make useful holographic interconnects, a mathematical model is required both to predict and analyse the effect of a grating on an incident wave. This chapter briefly discusses the various theories that exist to model the diffraction of waves by spatially periodic media in the far-field (Fraunhofer regime); pointing out which of these are applicable to volume HOEs and concentrating on the rigorous coupled-wave method [MOH81b]. For an analysis of Fresnel-type diffraction theory see references [NOP92, chapter 7] and [NOP94a].

3.1 Theoretical grating models

A spatially modulated periodic medium can be a dielectric, finitely or perfectly conducting material that is modulated in refractive index and/or absorption, either continuously (volume grating) or discontinuously (surface-relief grating). The principal models which have been applied to such media are:

- Fourier analysis methods
- Integral methods; and
- Differential methods

Which of these models is appropriate depends primarily upon the medium and form of modulation. In addition, some theories are approximate and others are

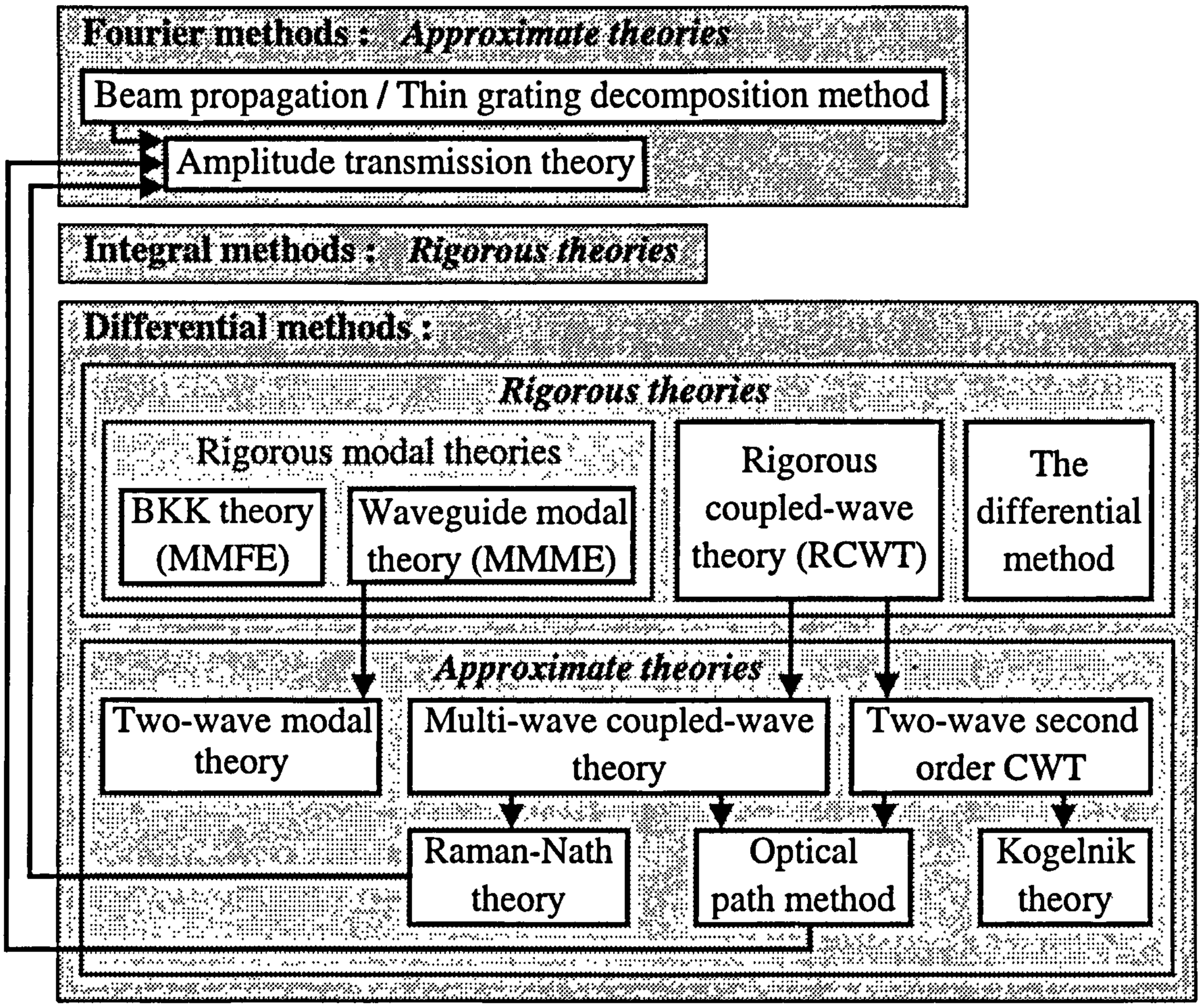


Figure 3.1: Schematic relationship between the most widely used surface-relief and volume diffraction grating models. The hierarchical structure illustrates how the volume grating theories discussed by Gaylord & Moharam [GAY85] fit into the general classifications discussed in this chapter.

rigorous solutions of the electromagnetic wave propagation inside these media. Rigorous theories represent a complete solution of Maxwell's equations for TE, TM or arbitrarily polarised radiation both inside and out of the modulated medium. Approximate theories make certain assumptions about the diffractive structure that simplify the mathematical model, most of which are related to the optical thickness of the grating (see appendix A). All of the models require numerical solutions; rigorous theories being more computationally intensive than their approximate counterparts. Several overviews compare and contrast these theories in detail [PET80, RUS81, MAY84, GAY85]. The diagram in figure 3.1 illustrates the basic relationship between the most widely used grating models and a brief sum-

mary of the salient points follows.

3.1.1 Fourier analysis methods

Fourier methods are *approximate* theories which are most widely used to model 'thin' volume and surface relief gratings, but can be adapted to model thick volume HOEs. These methods fall into two broad categories:

Amplitude transmission theory

This is generally considered a 'thin' grating theory and is widely used in the design of binary surface-relief structures [DAM71, VAS89, WYR91, VAS92, MIL93]. The method is equally applicable to absorption and phase modulated gratings; involving the calculation of the energy spectrum of the diffracted waves from the Fourier transform of the grating description. Multiple diffraction orders are included and complicated grating shapes are allowed, but small modulation and incident radiation parallel to the fringes (paraxial domain) are assumed. These approximations make it unsuitable for the modelling of 'thick' volume phase gratings.

Beam propagation method

Is widely used in the modelling of volume HOEs [ALF75b, ALF76, STO86, ICH93]. The volume grating is decomposed into a series of thin gratings within which the amplitude transmission theory (ATT) can be utilised. When applied specifically to volume gratings it is often referred to as the Thin Grating Decomposition Method (TGDM), [ALF75b, ALF76]. The use of ATT allows for arbitrary grating shapes but also facilitates the modelling of multiple superposed gratings [ICH93].

3.1.2 Integral methods

A detailed analysis of integral methods is presented by Maystre [MAY80]. They are rigorous and involve the solution of Maxwell's equations inside the modulated region by integral methods, (not to be confused with the numerical integration used in some solutions of differential methods). This is very complicated both theoretically and

in the numerical solution. They are not suitable for graded-index gratings such as volume HOEs, but are ideal for analysing all forms of surface-relief structures.

3.1.3 Differential methods

These are the most widely used and flexible theories in use [VIN80, GAY85]. They are rigorous and involve the solution of Maxwell's equations inside the modulated region by differential methods. The explicitly titled 'Differential methods', reviewed by Vincent [VIN80], are grouped for convenience with other methods in which Maxwell's equations are solved by differential techniques. Vincent's method is not described here because it is not usually applied to volume gratings. The two most commonly implemented methods are:

- Modal
- Coupled-wave

They can be applied to all forms of grating and simplified, (by making assumptions about the grating) to approximate theories. The limitations of differential with respect to integral methods (for surface-relief gratings) only arise in the stability and speed of convergence of the numerical solution.

Modal methods

Rigorous modal methods can be split into two classes [LI93a], those where the electromagnetic field inside the grating is represented by either a modal or Fourier expansion:

1. Modal method with a modal expansion (MMME). This is the classical waveguide mode theory in which the field inside each individual groove (for lamellar gratings) or between fringe planes (volume gratings) is expressed as a waveguide mode. The total field inside the modulated region is then expressed as a sum of all possible modes over an infinitely periodic medium.

2. Modal method with a Fourier expansion (MMFE). The total field inside the modulated region is expressed as a summation of modes of the entire diffracting

structure. Each mode being expressed as a Fourier sum and the modulation being represented as a Fourier series.

Methods utilising a Fourier expansion of the permittivity profile [VAS91, LI93a] are often referred to as BKK theories, after Burckhardt, Kaspar and Knop [BUR66, KAS73, KNO78]. Both methods can treat a grating structure as either a slab (volume), where the modulation is allowed to vary throughout the medium, or as a series of lamellar gratings (multi-layer model similar to a TGDM) in which the modulation is invariant through the depth of each layer.

Volume holograms are considered by Burckhardt [BUR66] and Kaspar [KAS73] applying the BKK method to gratings with unslanted fringes, whereas Chu and Kong [CHU77a] apply the waveguide method to slanted fringes. However, modal methods are more commonly applied to the design and analysis of surface-relief gratings, some of these utilise a modal expansion [MAY72, BOT81, SHE82, VAS92, LI93b, MIL93, LI94] and others a Fourier expansion [KNO78, VAS91, VAS92, NOP92, LI93a]. Which method is most suitable depends upon the type of grating under consideration; there are trade-offs concerning the computational complexity, numerical stability and speed of convergence of the model. An analysis of these issues has been carried out by Li & Haggans [LI93a] and Miller et al. [MIL94].

A general slanted grating BKK method, treating the medium as a slab, has not been put forward to date. Such a method would be desirable, since the BKK approach can be formulated to be less computationally intensive and numerically more stable than the rigorous coupled-wave method [VAS91].

An *approximate* modal theory analogous to the MMME and applied to volume holograms is the two-wave modal theory [BER67, LED80]. Only the zero and first-order diffracted waves are retained in this analysis.

Coupled-wave methods

The rigorous CWM, (sometimes confusingly called the coupled-mode method), is applicable to both volume [MOH81b, MOH83b] and surface-relief gratings [MOH82b]. It differs from the rigorous modal methods only in the representation of the elec-

tromagnetic field inside the grating. The two methods have been proven to be analytically and numerically equivalent [MAG78b, GAY85]. The field expansion is in terms of plane waves (or space-harmonics) each of which does not satisfy the wave equation but the summation of which does. This is in contrast to the modal methods in which each wave mode is itself a solution to the wave equation.

In both rigorous MMs and CWMs boundary diffraction, multiple diffraction orders (in transmission and reflection) and all differential terms are included. This makes them relatively complicated and numerically intensive to solve when compared to simpler approximate methods.

There are many *approximate* CWMs, each of which omits one or more of the three rigorous criteria above:

- Multi-wave coupled-wave theory; in which only first-order differential terms are retained and boundary diffraction effects are ignored [MAG77, RED89a].

- Two-wave second-order coupled-wave theory; in which only the zero- and first-order diffracted waves are retained (similar to two-wave modal theory) [KOW76, KON77].

- Two-wave coupled-wave theory (often called Kogelnik theory); in which only first-order differential terms are retained, boundary effects are ignored and only zero and first-order diffracted waves are included [PHA56, KOG69].

- Optical path method; which has similar assumptions to multi-wave theory added to that of a small modulation [RAM35, KLE67, SYM80].

- Raman-Nath theory; which again has similar assumptions to multi-wave theory but the Bragg condition for all diffracted orders are simultaneously satisfied [RAM36, MAG78a].

Coupled-wave methods are preferred by many researchers working on volume HOEs because of the intuitive appeal of the field expansion in the modulated region. This leads to the phase matched forward and backward-diffracted plane waves that are so obvious experimentally, (see section 3.2), analogous to the optical recording

situation. However, the extra computational complexity of the RCWM over the modal methods has led, in many cases, to the latter being preferred for the design and analysis of surface-relief gratings.

3.1.4 Applicability of approximate theories

The applicability of an approximate theory is principally dependent upon whether the grating is optically 'thick' or 'thin' (see appendix A) and whether it is predominantly a transmission or reflection grating. This second point is determined by what is known as 'the slowly varying amplitude approximation'. The essence of this approximation is that if the coupling of waves in the modulated region is slow only first-order derivatives of the electromagnetic field need to be maintained in the analysis. This assumption is not valid for reflection gratings, (see appendix A).

Approximate theories are very useful because of their relative numerical simplicity and ease of solution. However, rigorous treatment becomes essential when the dimensions of the grating modulation approach the wavelength of the incident EM radiation, (this is particularly true for surface-relief structures).

The relationship and equivalence between the various rigorous and approximate theories is discussed by Magnusson, Moharam and Gaylord in several papers [MAG78b, MOH81b, GAY85].

3.1.5 Assumptions in rigorous theories

Even rigorous theories make one assumption: that the diffracting structure is infinite, generally implying illumination by an infinite uniform plane wave. Some work has been carried out on diffraction from grating structures by finite non-uniform profile beams of light [CHU77b, MOH80b, HAR92]. This assumption is, however, generally valid; as long as a significant number of grating periods are illuminated the grating will behave as an infinite periodic structure.

Most theories are two-dimensional and only applicable to planar gratings. This two-dimensionality allows the problem to be broken down into simpler single polarisation solutions, either TE or TM (e.g. see [MOH81b] and [MOH83b] respectively).

However, some vector three-dimensional theories have been developed for incident radiation of any polarisation at any angle in three-dimensional space (often referred to as conical incidence) [MOH83c, GLY87, GLY89].

3.1.6 Accounting for non-linear volume media

The theories discussed above deal with ideal linear recording media. Extensions to the basic theories have been made by researchers who characterise recording media and require a complete theoretical fit with experimental data to explain both the non-uniformities of holographic fringes, (in orientation and modulation), and those of bulk refractive index. In DCG these non-uniformities are the direct result of wet processing and absorption variation with depth in the emulsion during exposure. Some theories are more easily modified to incorporate such an analysis than others - in particular the thin grating decomposition method.

DCG has been characterised in this way [NEW87, RED89a], the influences of processing non-uniformities are well understood and will not be considered here. The requirements of the model in this thesis are to predict the behaviour of a wide range of volume phase gratings which are to be utilised in both transmission and reflection interconnects. The most general and easily implementable of these theories is the rigorous coupled-wave method (RCWM) of Moharam & Gaylord [MOH81b]. This is applicable to TE (or $H-$mode) polarised incident light and is not generally held true for conformal reflection gratings (see section 3.8.3), but is otherwise completely accurate, (the theory has been applied to $E-$mode polarisation [MOH83b] and generalised to conical incidence [GLY87]).

The theory presented here is a new generalisation of Moharam & Gaylords original paper [MOH81b], taking into account absorption modulation and harmonic gratings. For simplicity, most of Moharam & Gaylords notation is maintained, any deviation from this is required in order to generalise the theory, still further, to model multiple gratings in the same volume, (covered in chapter 5).

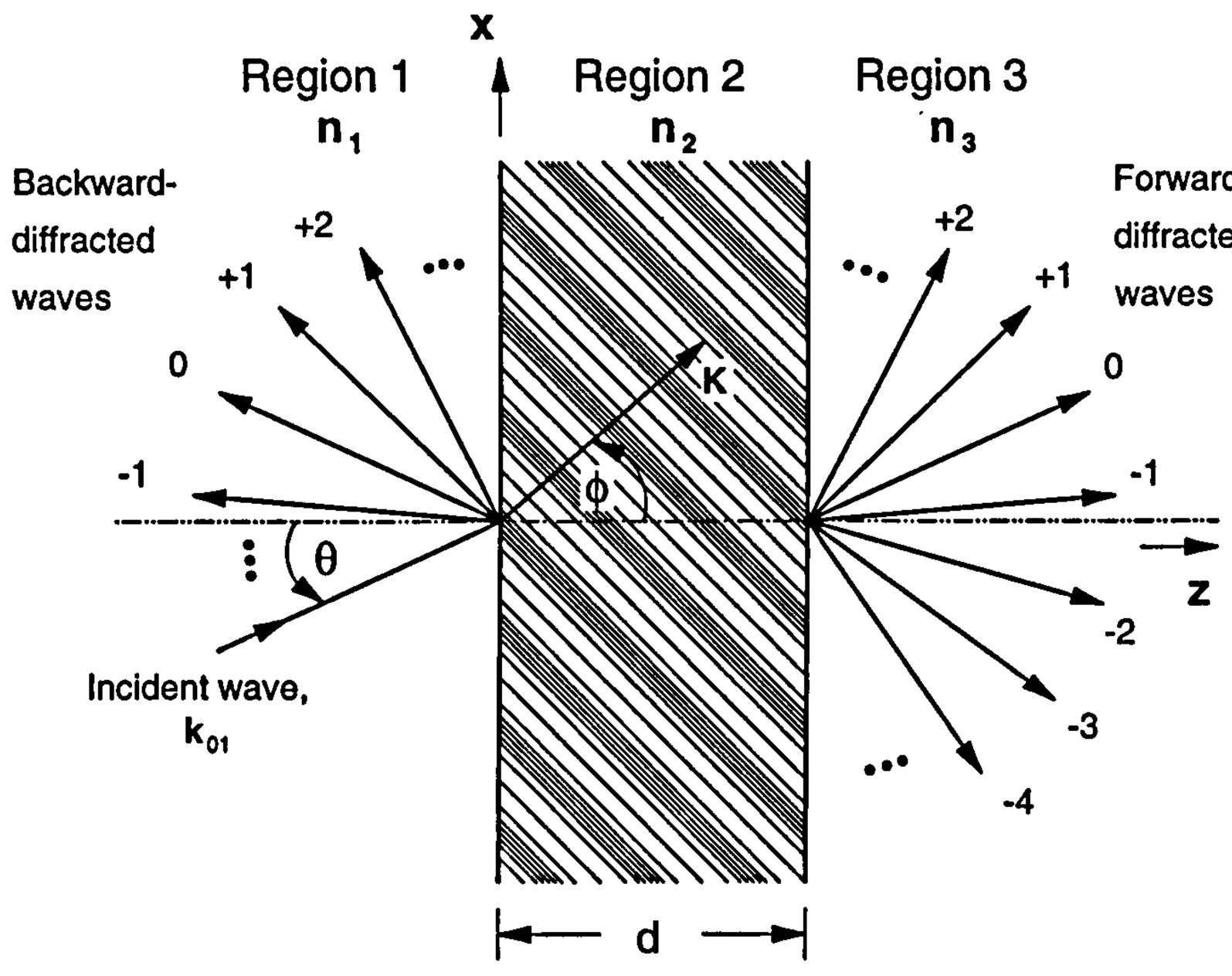


Figure 3.2: Hologram geometry for planar grating diffraction models.

3.2 The rigorous coupled-wave method

Optically recorded holograms can be fabricated using completely arbitrary wavefronts. This flexibility facilitates the implementation of highly complex optical interconnects. Most volume hologram theories, however, restrict themselves to diffracting structures formed by the interference of infinite plane-waves. Rigorous coupled-wave theory is similarly restrictive and is further constrained to gratings formed by the interference of only two waves; a reference and an object wave. Both of these waves are assumed to be incident in the $x-z$ plane, recording a refractive index and/or absorption grating modulated in this plane. The reconstruction wave is also assumed to lie in this plane, making an angle θ with respect to the z—axis, as shown in figure 3.2. This wavefront is considered to be an infinite homogeneous plane-wave of unit amplitude, plane-polarised normally to the plane of incidence (TE polarisation). The propagation of such a reconstruction wavefront, through the hologram

(region 2), is described by the scalar wave (Helmholtz) equation:

$$\nabla^2 E_2 + k^2 E_2 = 0 \,, \tag{3.1}$$

where

$$k^2 = \left(\frac{\omega^2}{c^2}\right) n^2 - 2i \left(\frac{\omega}{c}\right) n\alpha \,, \tag{3.2}$$

E_2 is the complex amplitude of the y–component of the electric field in region 2, c is the speed of light in a vacuum and ω is the angular frequency of the replay wavefront. The refractive index, n, and absorption coefficient, α, in the modulated region of a material with a linear response are given by:

$$\begin{aligned} n &= n_2 + \Delta n_2 \text{ ; and} \\ \alpha &= \alpha_2 + \Delta\alpha_2 \,, \end{aligned} \tag{3.3}$$

where n_2 and α_2 are the average refractive index and absorption within the hologram. The terms Δn_2 and $\Delta\alpha_2$ are the spatially varying refractive index and absorption modulations which totally describe the fringe profiles within the hologram. Moharam & Gaylord [MOH81b] originally assumed a pure sinusoidal phase grating ($\alpha = 0$) and described the refractive index, n, in the modulated region in terms of the relative permittivity, ϵ (equal to n^2), where $\epsilon(x,z) = \epsilon_2 + \Delta\epsilon \cos\left[K\left(x\sin\phi + z\cos\phi\right)\right]$. This would simplify equation 3.2 to $k = {}^{2\pi\epsilon}/_{\lambda}$. In this work these descriptions have been extended to cover gratings with arbitrary profile (including harmonics) and to incorporate absorption modulation:

$$\begin{aligned} \Delta n_2 &= \sum_{h=1}^{N_h} [\Delta n_{ch} \cos(h\mathbf{K}.\mathbf{r}) + \Delta n_{sh} \sin(h\mathbf{K}.\mathbf{r})] \text{ ; and} \\ \Delta\alpha_2 &= \sum_{h=1}^{N_h} [\Delta\alpha_{ch} \cos(h\mathbf{K}.\mathbf{r}) + \Delta\alpha_{sh} \sin(h\mathbf{K}.\mathbf{r})] \end{aligned} \tag{3.4}$$

Where $\mathbf{K}.\mathbf{r} = K(x\sin\phi + z\cos\phi)$, N_h is the number of Fourier harmonics under consideration, $h\mathbf{K}$ is the wave vector of the h:th harmonic grating of the fundamental grating $\mathbf{K}$ and Δn_{ch}, $\Delta\alpha_{ch}$, Δn_{sh} and $\Delta\alpha_{sh}$ are the Fourier cosinusoidal and sinu-

soidal amplitudes, of refractive index and absorption modulation for harmonic h. These harmonic amplitudes can be deduced by Fourier analysis.

3.2.1 The vector Floquet condition

In order to solve the Helmholtz equation 3.1 an expression for the total field inside the grating region, E_2, must be considered. Gaylord & Moharam [GAY85] show that a number of alternative (but equivalent) field representations exist. The fundamental coupled-wave expansion assumes that the modulation of the grating is periodic and infinite in extent along the $x-$axis. Under such conditions E_2 can be expanded as a Fourier summation of inhomogeneous plane-waves propagating in the $x-$direction [MAY84]:

$$E_2 = \sum_{m=-\infty}^{+\infty} A_m(z) \exp[-i(k_{02x} + mK_x)x] \tag{3.5}$$

where A_m is the normalised amplitude of the m:th inhomogeneous plane-wave component, k_{02x} is the $x-$component of the wavevector of the refracted incident wave and K_x is the $x-$component of the grating vector. This expansion leads to a non-constant coefficient set of differential equations, when inserted into the wave equation, making a direct solution difficult. However, a set of constant coefficient linear differential equations results from the definition of a new function [GAY85]:

$$S_m(z) = A_m(z) \exp[+i(k_{02z} + mK_z)z] \tag{3.6}$$

where k_{02z} and K_z are the $z-$components of the refracted incident wavevector and grating vector respectively. The electric field inside the grating may now be expressed in a vector form as:

$$E_2 = \sum_{m=-\infty}^{+\infty} S_m(z) \exp[-i\mathbf{k}_m.\mathbf{r}] \tag{3.7}$$

where

$$\mathbf{k}_m = \mathbf{k}_{02} + m\mathbf{K} \tag{3.8}$$

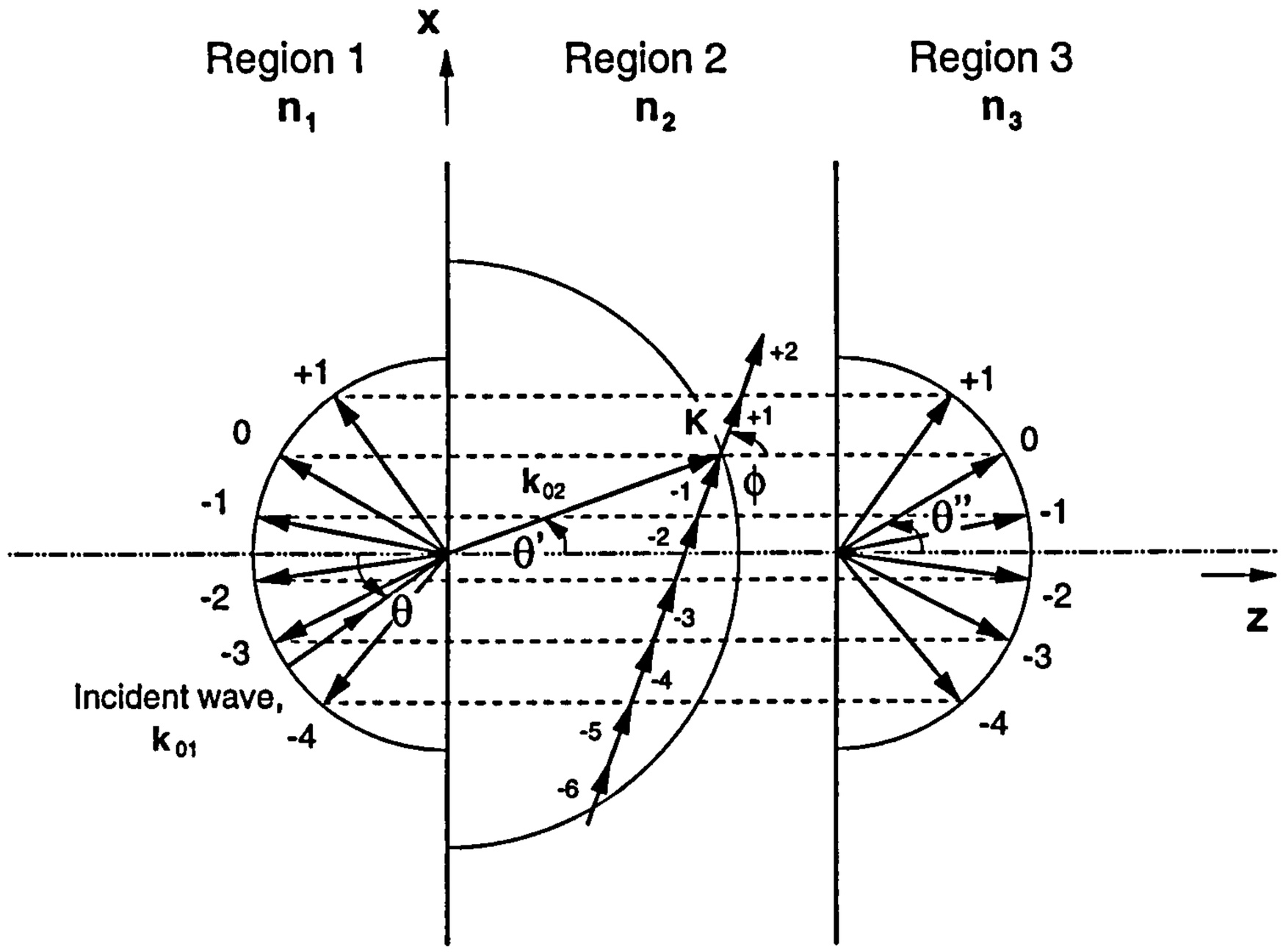


Figure 3.3: Wavevectors for a general slanted grating from the vector Floquet condition, (the situation depicted assumes $n_1 = n_3$).

The vector equation 3.8 is known as the vector Floquet condition; it defines the diffracted direction of all the space-harmonics in the modulated region. Figure 3.3 shows the wavevectors, both inside and outside the modulated region, arising from this condition for a general slanted grating. The space-harmonic components inside the modulated region are phase matched with the propagating forward and backward-diffracted orders, (shown by the dashed lines). There are six propagating diffracted orders shown in figure 3.3, ($-4 \leq m \leq +1$), all other orders are evanescent ($m \leq -5$ and $m \geq +2$).

3.2.2 Evanescent waves in dielectric media

Evanescent or surface waves travel along the boundary between two dielectric media. The energy in these waves circulates back and forth across the interface, resulting in an average of zero net power flow into the less dense medium, thereby conserving energy [HEC87, page 107], (time-averaged Poynting vector is zero). The amplitude and phase fronts, of these waves, are not co-planar and their amplitude decays

exponentially with depth outside of the modulated region.

Evanescent diffraction orders behave in the same way as propagating orders within the modulated region - they are inhomogeneous and energy couples between all of the space-harmonics. Only at the dielectric boundary do they exist as surface waves. If they are not included in the analysis an essential part of the coupling mechanism is ignored. In this way they affect the accuracy of the analysis and convergence of the numerical solution to an absolute stable value, (discussed in section 3.4).

An effective demonstration of the importance of evanescent waves to grating fabrication and analysis is illustrated in the copying of CGHs into volume media. This is usually done using an interferometric recording arrangement (see figure 2.6) in which one arm of the system images a CGH amplitude mask, via a $4f$ optical system, onto the holographic plane [ROB93a, chapter 4]. Using such an arrangement, (even with large NA lenses), it is impossible to exactly replicate the field generated by the grating because evanescent orders do not propagate and are therefore not imaged. The volume copy will consequently not replay quite as accurately as expected because it does not represent a faithful duplication of the original. Further evidence of this can be seen in the Talbot self-imaging work of Noponen [NOP92, chapter 5].

3.2.3 Derivation of the coupled-wave equations

The Helmholtz equation can be solved in regions 1 and 3 (where the refractive index is constant) as a summation of plane-waves [RAY07]. Only a discrete set of these waves (determined by the vector Floquet condition) are allowed if region 2 is a periodic grating. Consequently, when a unit-amplitude plane-wave is incident in region 1, the normalised electric field amplitude in this region may be expressed as a summation of homogeneous plane-waves of the form:

$$E_1 = \exp[-i(\beta_0 x + \xi_{10} z)] + \sum_{m=-\infty}^{+\infty} R_m \exp[-i(\beta_m x - \xi_{lm} z)] \qquad (3.9)$$

where R_m is the normalised amplitude of the m:th reflected diffraction order, $\beta_m = k_{01} \sin\theta + mK \sin\phi$, for any integer m, where m is the wave index, $\xi_{lm} = k_{0l}^2 - \beta_m^2$ for

$l = 1, 3$ and $k_{0l} = {}^{2\pi n_l}/_{\lambda}$ for $l = 1, 2$ and 3. Here, l is the region index. Similarly, the normalised electric field amplitude in region 3 is given by the Rayleigh expansion:

$$E_3 = \sum_{m=-\infty}^{+\infty} T_m \exp[-i(\beta_m x + \xi_{3m}(z-d))] \tag{3.10}$$

where T_m is the normalised amplitude of the m:th transmitted diffraction order. The electric field in the modulated region (from equation 3.7) can be re-written as:

$$E_2 = \sum_{m=-\infty}^{+\infty} S_m(z) \exp[-i(\beta_m x + \xi_{2m} z)] \tag{3.11}$$

where $\xi_{2m} = k_{02}\cos\theta' + mK\cos\phi$, θ' is the refracted angle (inside the modulated region) of the incident wave and S_m is the normalised amplitude of the m:th wave field anywhere in region 2. The β and ξ terms are the $x-$ and $z-$components of the diffracted wavevectors derived from the vector Floquet condition, equation 3.8. These components form the basis of the wavevector expansion that makes the problem solvable using equation 3.11.

The wave equation in the modulated region is solved by substituting equations 3.3, 3.4 and 3.11 into equation 3.1, (ignoring terms in Δn_2^2, $\Delta\alpha_2^2$ and $\Delta n_2 \Delta\alpha_2$ which are negligible), and results in a set of infinite coupled wave equations:

$$\begin{aligned}\frac{d^2 S_m(u)}{du^2} &= \frac{2\xi_{2m}}{\kappa}\frac{dS_m(u)}{du} + \frac{(k_{02}^2 - \xi_{2m}^2 - \beta_m^2 - 2ik_{02}\alpha_2)}{\kappa^2} S_m(u) \\ &\quad + \sum_{h=1}^{N_h}\left(\frac{2\pi}{\kappa^2\lambda}\right)(D_h S_{m+h}(u) + E_h S_{m-h}(u))\end{aligned} \tag{3.12}$$

where

$$\begin{aligned} D_h &= k_{02}A_h - i(n_2 B_h + \alpha_2 A_h)\,, \\ E_h &= k_{02}A_h^* - i(n_2 B_h^* + \alpha_2 A_h^*)\,, \\ A_h &= \Delta n_{ch} - i\Delta n_{sh}\ ; \text{ and} \\ B_h &= \Delta\alpha_{ch} - i\Delta\alpha_{sh} \end{aligned} \tag{3.13}$$

The reduced variable u is introduced to simplify the complex analysis at this stage,

and the new parameters are given by $S_m(u) = S_m(z)$ and $u = {}^{(i\pi z)}/_{(2\lambda n_2)} = i\kappa z$, (note that κ is not the same as [MOH81b]).

The derivation of equation 3.12 is a straight-forward mathematical exercise. The only subtle point of note is that the coupling of space-harmonics to neighbouring waves is introduced as a direct consequence of the periodic terms in the expressions for refractive index and absorption. A detailed mathematical treatment of the coupled-wave formulation for TE polarisation can be found in Solymar & Cooke [SOL81, pages 115-117].

3.3 Method of solution

Unlike the rigorous method of Vincent [VIN80], the coupled-wave equations derived using this RCWM cannot be solved using Runge-Kutta techniques.

The solution to equation 3.12 is not a trivial one, but it can be solved numerically by a state equation approach, (see section 3.4.1). The state equation approach redefines this equation in a matrix form such as:

$$\begin{bmatrix} \dot{\mathbf{S}} \\ \ddot{\mathbf{S}} \end{bmatrix} = \begin{bmatrix} \\ b_{rs} \\ \\ \end{bmatrix} \begin{bmatrix} \mathbf{S} \\ \dot{\mathbf{S}} \end{bmatrix} \tag{3.14}$$

Where $\mathbf{S}$, $\dot{\mathbf{S}}$ and $\ddot{\mathbf{S}}$ indicate column vectors with components S_m, $\frac{dS_m}{du}$ and $\frac{d^2S_m}{du^2}$ respectively. The vector array b_{rs} is the coefficient matrix specified from equation 3.12. Equation 3.14 corresponds to an unforced state equation in the state-space description of linear systems and has a relatively straightforward solution. The solution is obtainable in terms of the eigenvalues and eigenvectors of the coefficient matrix b_{rs}, and is a summation of plane waves thus:

$$S_m(u) = \sum_{\nu=-\infty}^{+\infty} C_\nu w_{m\nu} \exp(q_\nu u) \tag{3.15}$$

where q_ν is the ν:th eigenvalue and $w_{m\nu}$ is the ν:th element of the row in the matrix $[w]$ composed of the eigenvectors corresponding to the m:th wave. The coefficients C_ν are unknown constants to be determined by solving the boundary

value problem i.e. by matching the tangential electric and magnetic fields at the two boundaries of the holographic medium.

3.3.1 Boundary conditions

The tangential electric fields at the boundaries of regions 1, 2 and 3 can be matched directly using equations 3.9, 3.10 and 3.11. The magnetic field, however, must be obtained from the same equations by applying the Maxwellian relation $\nabla \times \mathbf{E} = -\mu_0 \frac{d\mathbf{H}}{dt}$ before field matching can be performed. Matching the fields at the boundaries gives the following infinite set of equations. Tangential E at $z = 0$ gives:

$$R_m + \delta_{m0} = \sum_{\nu=-\infty}^{+\infty} C_\nu w_{m\nu} \tag{3.16}$$

tangential H at $z = 0$ gives:

$$\xi_{1m}(R_m - \delta_{m0}) = \sum_{\nu=-\infty}^{+\infty} C_\nu w_{m\nu}(q_\nu \kappa - \xi_{2m}) \tag{3.17}$$

tangential E at $z = d$ gives:

$$T_m = \sum_{\nu=-\infty}^{+\infty} C_\nu w_{m\nu} \exp[i(q_\nu \kappa - \xi_{2m})d] \tag{3.18}$$

and tangential H at $z = d$ gives:

$$-\xi_{3m} T_m = \sum_{\nu=-\infty}^{+\infty} C_\nu w_{m\nu}(q_\nu \kappa - \xi_{2m}) \exp[i(q_\nu \kappa - \xi_{2m})d] \,, \tag{3.19}$$

where δ_{m0} is the Kronecker delta function. Eliminating transmitted and reflected wave amplitudes, T_m and R_m, from these equations gives a set of simultaneous equations solvable for the coefficients C_ν:

$$\sum_{\nu=-\infty}^{+\infty} C_\nu w_{m\nu}(\xi_{2m} - q_\nu \kappa + \xi_{1m}) = 2\delta_{m0}\xi_{1m} \tag{3.20}$$

$$\sum_{\nu=-\infty}^{+\infty} C_\nu w_{m\nu}(\xi_{2m} - q_\nu \kappa - \xi_{3m}) \exp[i(q_\nu \kappa - \xi_{2m})d] = 0 \tag{3.21}$$

Substitution of the constants, C_ν, into equations 3.16 and 3.18 gives the reflected

and transmitted wave amplitudes from which the propagating diffraction order efficiencies can be calculated.

3.3.2 Diffraction efficiency

The sum of the diffraction efficiencies of all propagating waves must, by energy conservation, be unity:

$$\sum_{m=-\infty}^{+\infty} (\eta_{1m} + \eta_{3m}) = 1 \tag{3.22}$$

where η_{1m} and η_{3m} are the diffraction efficiencies, in regions 1 and 3, of wave m. These diffraction efficiencies are given by the $z-$component of the Poynting vector:

$$\eta_{1m} = \Re\left(\frac{\xi_{1m}}{\xi_{10}}\right) R_m R_m^* \tag{3.23}$$

and

$$\eta_{3m} = \Re\left(\frac{\xi_{3m}}{\xi_{10}}\right) T_m T_m^* \,, \tag{3.24}$$

where $\Re$ denotes the real part of the complex expression for the propagation constants ξ of the input wave and the diffracted orders. The experimental definition of diffraction efficiency is the diffracted power of a particular propagating order divided by the power in the input wave. The propagating orders are spatially separated as they move away from the grating. This coupled-wave formulation has been derived for infinite plane waves but if the diffracted waves were infinite they would all be present at each point in space and interference effects would complicate the interpretation of diffraction efficiency. However, the finite width of a beam of light is very large in comparison to the wavelength of the light and so the plane wave model is accurate; a finite beam width also implies that the propagating diffraction orders spatially separate after the grating. To calculate the diffraction efficiencies for waves of infinite extent from this model requires the time-averaged Poynting vector to be determined for each propagating order, as explained by Russell [RUS84]. The equations 3.23 and 3.24 above define the diffraction efficiencies in regions 1 and 3 in terms of the spatially-averaged Poynting vectors of propagating orders, since experimentally diffraction orders become spatially separated for beams of finite width.

3.4 Numerical solution

As discussed in section 3.3, the coupled wave equation 3.12 is not analytically solvable and requires the application of numerical techniques to investigate it. This section elaborates upon the state equation solution.

Numerical analysis necessitates the truncation of the set of infinite equations into a finite set of space-harmonics in the modulated region. Truncation of the model affects the accuracy of the solution, the more space-harmonics that are included in the model the more accurate the result is. For 'thick' volume holograms there are very few propagating orders and the inclusion of a large number of evanescent waves does not significantly affect the overall result from an experimental point of view. This is fortunate because as will be discussed in section 3.5, this algorithm can become numerically unstable with significant numbers of evanescent waves [CHA93, CHA94, MOH95a, MOH95b]. Russell [RUS84] discusses the absolute accuracy and power conservation issues of truncated coupled-wave theories.

If the coupled-wave equations are truncated from an infinite set to $2N+1$ waves, where $-N \leq m \leq +N$, then the simultaneous equations in C_ν are truncated to $4N+1$ equations, where $-2N \leq \nu \leq +2N$. This truncates the vector array $b_{r,s}$ into four sub-matrices of $N \times N$ as illustrated in equation 3.29

3.4.1 State equation approach

This is an elegant solution and involves the splitting of the second-order differential equations into two sets of first-order differential equations, by defining the state variables:

$$S_{1,m} = S_m(u) \tag{3.25}$$

$$S_{2,m} = \frac{dS_m(u)}{du} = \dot{S}_m(u) \tag{3.26}$$

This leads to the two first-order constituent state equations:

$$\begin{aligned} \frac{dS_{1,m}}{du} &= \dot{S}_{1,m} = S_{2,m} \\ \frac{dS_{2,m}}{du} &= \dot{S}_{2,m} = a_h S_{1,m+h} + b_m S_{1,m} + a_h' S_{1,m-h} + c_m S_{2,m} \end{aligned} \tag{3.27}$$

where

$$\begin{aligned} a_h &= \left(\frac{2\pi}{\kappa^2\lambda}\right) D_h \,, \\ a_h^{'} &= \left(\frac{2\pi}{\kappa^2\lambda}\right) E_h \,, \\ b_m &= \frac{k_{02}^2 - \xi_{2m}^2 - \beta_m^2 - 2ik_{02}\alpha_2}{\kappa^2} \text{ ; and} \\ c_m &= \frac{2\xi_{2m}}{\kappa} \end{aligned} \tag{3.28}$$

Writing equations 3.27 as a single matrix state equation gives equation 3.14 which in an expanded form looks like equation 3.29.

$$\begin{bmatrix} \vdots \\ \dot{S}_{1,2} \\ \dot{S}_{1,1} \\ \dot{S}_{1,0} \\ \dot{S}_{1,-1} \\ \dot{S}_{1,-2} \\ \vdots \\ \dot{S}_{2,2} \\ \dot{S}_{2,1} \\ \dot{S}_{2,0} \\ \dot{S}_{2,-1} \\ \dot{S}_{2,-2} \\ \vdots \end{bmatrix} = \begin{bmatrix} & & & \vdots & & & & & \vdots & & & \\ & 0 & 0 & 0 & 0 & 0 & & 1 & 0 & 0 & 0 & 0 & \\ & 0 & 0 & 0 & 0 & 0 & & 0 & 1 & 0 & 0 & 0 & \\ \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots \\ & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 1 & 0 & \\ & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 1 & \\ & & & & \vdots & & & & & \vdots & & & \\ & b_2 & a_1 & a_2 & a_3 & a_4 & & c_2 & 0 & 0 & 0 & 0 & \\ & a_1^{'} & b_1 & a_1 & a_2 & a_3 & & 0 & c_1 & 0 & 0 & 0 & \\ \cdots & a_2^{'} & a_1^{'} & b_0 & a_1 & a_2 & \cdots & 0 & 0 & c_0 & 0 & 0 & \cdots \\ & a_3^{'} & a_2^{'} & a_1^{'} & b_{-1} & a_1 & & 0 & 0 & 0 & c_{-1} & 0 & \\ & a_4^{'} & a_3^{'} & a_2^{'} & a_1^{'} & b_{-2} & & 0 & 0 & 0 & 0 & c_{-2} & \\ & & & & \vdots & & & & & \vdots & & & \end{bmatrix} \begin{bmatrix} \vdots \\ S_{1,2} \\ S_{1,1} \\ S_{1,0} \\ S_{1,-1} \\ S_{1,-2} \\ \vdots \\ S_{2,2} \\ S_{2,1} \\ S_{2,0} \\ S_{2,-1} \\ S_{2,-2} \\ \vdots \end{bmatrix} \tag{3.29}$$

The state equation matrix 3.29 is solved to resolve its eigenvalues and vectors, these eigenvalues and vectors are then used to calculate the constants, C_ν, space-harmonic amplitudes R_m and T_m and the diffraction efficiencies of the propagating orders according to the method outlined in sections 3.3.1 and 3.3.2.

From the state equation 3.29 it is clear that the vector array $b_{r,s}$ can be viewed as four constituent sub-matrices. The sub-matrix defined in terms of the $a, a^{'}$ and b coefficients is directly analogous to the coupling matrix evident in other coupled-wave

$$\begin{bmatrix} \vdots \\ \dot{S}_{1,2} \\ \dot{S}_{1,1} \\ \dot{S}_{1,0} \\ \dot{S}_{1,-1} \\ \dot{S}_{1,-2} \\ \vdots \\ \dot{S}_{2,2} \\ \dot{S}_{2,1} \\ \dot{S}_{2,0} \\ \dot{S}_{2,-1} \\ \dot{S}_{2,-2} \\ \vdots \end{bmatrix} = \begin{bmatrix} & & & \vdots & & & & & \vdots & & & \\ & 0 & 0 & 0 & 0 & 0 & & 1 & 0 & 0 & 0 & 0 & \\ & 0 & 0 & 0 & 0 & 0 & & 0 & 1 & 0 & 0 & 0 & \\ \cdots & 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots \\ & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 1 & 0 & \\ & 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 1 & \\ & & & & \vdots & & & & & \vdots & & & \\ & b_2 & a & 0 & 0 & 0 & & c_2 & 0 & 0 & 0 & 0 & \\ & a & b_1 & a & 0 & 0 & & 0 & c_1 & 0 & 0 & 0 & \\ \cdots & 0 & a & b_0 & a & 0 & \cdots & 0 & 0 & c_0 & 0 & 0 & \cdots \\ & 0 & 0 & a & b_{-1} & a & & 0 & 0 & 0 & c_{-1} & 0 & \\ & 0 & 0 & 0 & a & b_{-2} & & 0 & 0 & 0 & 0 & c_{-2} & \\ & & & & \vdots & & & & & \vdots & & & \end{bmatrix} \begin{bmatrix} \vdots \\ S_{1,2} \\ S_{1,1} \\ S_{1,0} \\ S_{1,-1} \\ S_{1,-2} \\ \vdots \\ S_{2,2} \\ S_{2,1} \\ S_{2,0} \\ S_{2,-1} \\ S_{2,-2} \\ \vdots \end{bmatrix} \tag{3.30}$$

methods [RED89a, pages 69-73]. This analogy is exploited in chapter 5 in order to generalise the theory for fan-out gratings.

Equation 3.30 illustrates the original state equation matrix of Moharam & Gaylord [MOH81b] using the notation of this chapter. Note that in the absence of grating harmonics ($h = 1$) and absorption ($\alpha = 0$) then $a = (2n_2)^3 \Delta n_2$ and b_m & c_m are given by equation 3.28. It is evident from examining the coupling sub-matrices in equations 3.29 and 3.30 where the additional complexity of absorption modulation and an arbitrary grating profile (presented in this chapter) extend the analysis.

This model was implemented on a SUN Systems SPARC 10 model 51 in FORTRAN 77 using Numerical Algorithms Group [NAG93] routines, (to calculate eigenvalues and vectors and to solve the set of simultaneous equations).

3.5 Evanescent waves and numerical instabilities

The nature of the eigenvalue problem, fundamental to this model, can lead to numerical instabilities in the computational implementation of the theory. The problem arises in the solution of the simultaneous equations for the constants C_ν, in partic-

ular equation 3.21. This equation contains an exponential term, $\exp[i(q_\nu\kappa - \xi_{2m})d]$, where q_ν is the ν:th eigenvalue of the eigenvalue problem. The inclusion of evanescent space-harmonics in the analysis generates eigenvalues which are negative imaginary quantities. Under certain circumstances, such as relatively oblique incident angles or relatively thick gratings, these negative imaginary values can become very large indeed, such that combined with the exponential factor in this equation the arithmetic result exceeds the numerical accuracy of the computer; usually resulting in overflow. This is an unavoidable problem with this model that also exists in the BKK modal method, however in BKK theory the geometry can be reformulated to overcome this limitation [VAS92]. Chateau & Hugonin [CHA94] have developed and implemented an alternative rigorous coupled-wave formulation in which these numerical limitations are overcome, but it is significantly more complicated than this method. Moharam et al. [MOH95a, MOH95b] and Peng et al. [PEN95] have recently implemented formulations of the RCWM that are both efficient and numerically stable when applied to surface-relief gratings.

Provided one is aware that such problems can arise when a number (greater than three) of space-harmonics are included, the model is entirely satisfactory for analysing a whole range of volume holographic interconnects. The model encounters no problems for three space-harmonics ($+1, 0$ and -1 diffracted orders) over a whole range of differently specified gratings in both transmission and reflection; it can usually handle many more without instability. A slightly different formulation of the model is presented in chapter 5, in which it is possible to be selective about which space-harmonics are included in the analysis, enabling a larger number of orders without overflow. This selectivity is achieved at the expense of rigour and consequently a small degree of accuracy is lost.

3.6 Substrate modelling

This model deals with a relatively thin modulated dielectric film surrounded on both sides by a semi-infinite half-space of fixed dielectric constant (figure 3.2). DCG volume holograms do not exactly correspond to this situation: the gelatin layer requires

a rigid support (or substrate) to maintain its physical properties, (and to dissipate heat when used in certain applications). Thus a more accurate representation of a DCG HOE would include a glass substrate between the modulated film and one half-space.

The RCWM is not easily adapted to rigorously model this situation, although a formulation presented by Moharam & Gaylord for dielectric surface-relief gratings [MOH82b] should be able to account for this. Several other authors have included substrates into their model, notably:

- Vincent's differential method [VIN80, pages 115-117] is easily modified to include substrates.

- Chateau & Hugonin's RCWM [CHA93, CHA94] allows not only for substrates, but also dielectric coatings and DCG protection covers. However, in the volume HOE example considered in reference [CHA94], the substrate is modelled by assuming that the dielectric semi-infinite half-space after the hologram is glass, (region 3 of figure 3.2).

The approximation of Chateau [CHA94] is not necessarily the most appropriate simplified way of including a substrate into the analysis, particularly when analysing the zero-order transmission spectrum of an HOE, (this order contains all of the diffraction information). This is because orders are coupled out of the interaction region (propagate into region 3) that should be evanescent in the real situation. An alternative approach, that is particularly useful when analysing the transmission spectra of a reflection or transmission grating, is to model region 1 as the substrate i.e. a semi-infinite half-space of refractive index equal to that of the substrate [TUR94].

This approach is not exact, but since the DCG films used in this thesis have a thickness very much less than the substrate thickness (649F and Pilkingtons plates have a relative thickness between 60 and 80) it is a good approximation. Without the inclusion of a substrate, thin film Fabry-Perot effects are evident in the analysis that are not present in the experimental grating-substrate combination.

3.7 Exposure and saturation response

Although not fundamental to the rigorous coupled-wave model, it is important when designing HOEs to be able to relate exposure energy to the expected refractive index modulation. This will become apparent in the experimental considerations of chapter 4 (sections 4.5 and 4.6). A description is included here in order to explain the necessity of maintaining harmonic grating parameters in the theory, which have a bearing on the experimental results presented in section 3.8.

According to the simple theory of Kogelnik, (presented in sections 2.1 and 2.2), a transmission volume phase hologram replayed at Bragg incidence should have an exposure response that follows a $\sin^2$ function. This is only true, however, for an ideal linear recording medium in which the refractive index modulation is directly proportional to the exposure energy. Practical holographic recording materials only exhibit a linear response over a low exposure range, (usually up to the first diffraction maximum for a transmission grating). It has been well documented [CHA79, CHA80, NEW87, RED89a, ROB93b] that with increasing exposure the efficiency deviates from the Kogelnik model, due to saturation of the refractive index modulation. This is predictable from the discussion of section 2.5, in that the physical change that brings about the index modulation must have a finite limit.

Saturation of the refractive index modulation leads to a non-sinusoidal volume grating profile that must contain grating frequency components other than the fundamental. These must all be harmonics of the fundamental, (see equation 3.4). Several authors have suggested ways of modelling this saturation [CHA79, BLA90, NEW87], the method used here is that of Chang & Leonard [CHA79].

An exponential saturation of the index is assumed, leading to a description of the modulation, $\Delta n(x)$, in the direction of the grating vector of the form

$$\Delta n(x) = \Delta n_{max}\left[1 - \exp\left(\frac{-\beta E}{\Delta n_{max}}(1 + \cos Kx)\right)\right], \tag{3.31}$$

where Δn_{max} is the maximum available index modulation, β represents the photosensitivity of the medium (and is strongly wavelength dependent) and E is the ex-

posure energy. Fourier analysis of this saturated function (for details see [ROB93b, page 42]) results in the index modulation amplitude for each harmonic grating considered. The extended RCWM described in section 3.2 allows this full grating description to be included.

The parameters Δn_{max} and β determine the exposure response of a DCG hologram (for typical values see section 2.5.4). They are dependent on a wide variety of factors and in order to obtain reproducible results it is imperative that accurate control of all pre-processing, recording and processing parameters is maintained.

3.8 Grating response

To illustrate the power and versatility of the model the following sections present a theoretical fit to experimental data of a typical transmission and reflection grating. The zero-order transmitted power is shown in all cases because it contains information on all diffracted directions. The transmission grating is analysed in terms of an angular spectrum and the reflection grating a wavelength spectrum because they are more sensitive in this configuration, whereas only limited information could be obtained the other way around. Beforehand, however, some modifications to the model are discussed that are necessary to account for the processing of gelatin layers.

3.8.1 Processing changes

The processing of DCG emulsions, as discussed in chapter 2, results in two key unavoidable changes to the physical characteristics of the gelatin:

- **Bulk refractive index change.** Sensitised gelatin (before exposure) has a larger refractive index than exposed and processed gelatin. The size of this change varies depending upon sensitiser concentration, exposure energy and processing [NEW87]. Typically a sensitised plate has an index of 1.6 ± 0.02 and a processed plate 1.48 ± 0.02.

- **Emulsion thickness change.** A sensitised gelatin layer will have a thickness larger than an unsensitised layer and this thickness will change during the wet

processing of a DCG hologram. The thickness change is dependent upon the original gelatin thickness, bias hardness, exposure energy and fringe angle of the grating with respect to the gelatin surface [CHA79, NEW87]. Transmission gratings usually incur a net emulsion shrinkage and reflection gratings a net swelling.

These changes between the recording and replay of a hologram affect the Bragg characteristics of the grating. From the Bragg condition, (equations 2.3 and 2.4), it can be seen that these changes will alter the Bragg angle and the wavelength at which the grating is most efficient. In addition to the obvious shift in fringe spacing a change in emulsion thickness will alter the angle of the fringes and hence the grating vector [HAR84, pages 61-62].

Fortunately the changes in refractive index and thickness tend to counteract one another, usually resulting in small deviations in the Bragg characteristics. These can be recovered to the design specification with the post-process tuning described in section 2.8.

The implementation of the rigorous model presented here incorporates these refractive index, thickness and grating vector modifications.

3.8.2 Transmission grating analysis

The results presented in this section are for a typical volume grating, based on Kodak 649F gelatin, recorded and replayed at 514.5nm. The hologram is the same as that analysed by Redmond [RED89a] and consequently allows for a direct comparison between the multi-wave theory implemented by him (which incorporates grating non-uniformities) and the theory discussed above, based on the RCWM. The grating was recorded with beams at 0° and 30° and the experimental data has not been corrected for specular reflection or absorption losses.

Multi-wave model

Figure 3.4 represents the results of the multi-wave coupled-wave theory of Redmond. It includes second harmonics (clearly apparent as a minima at about 16°)

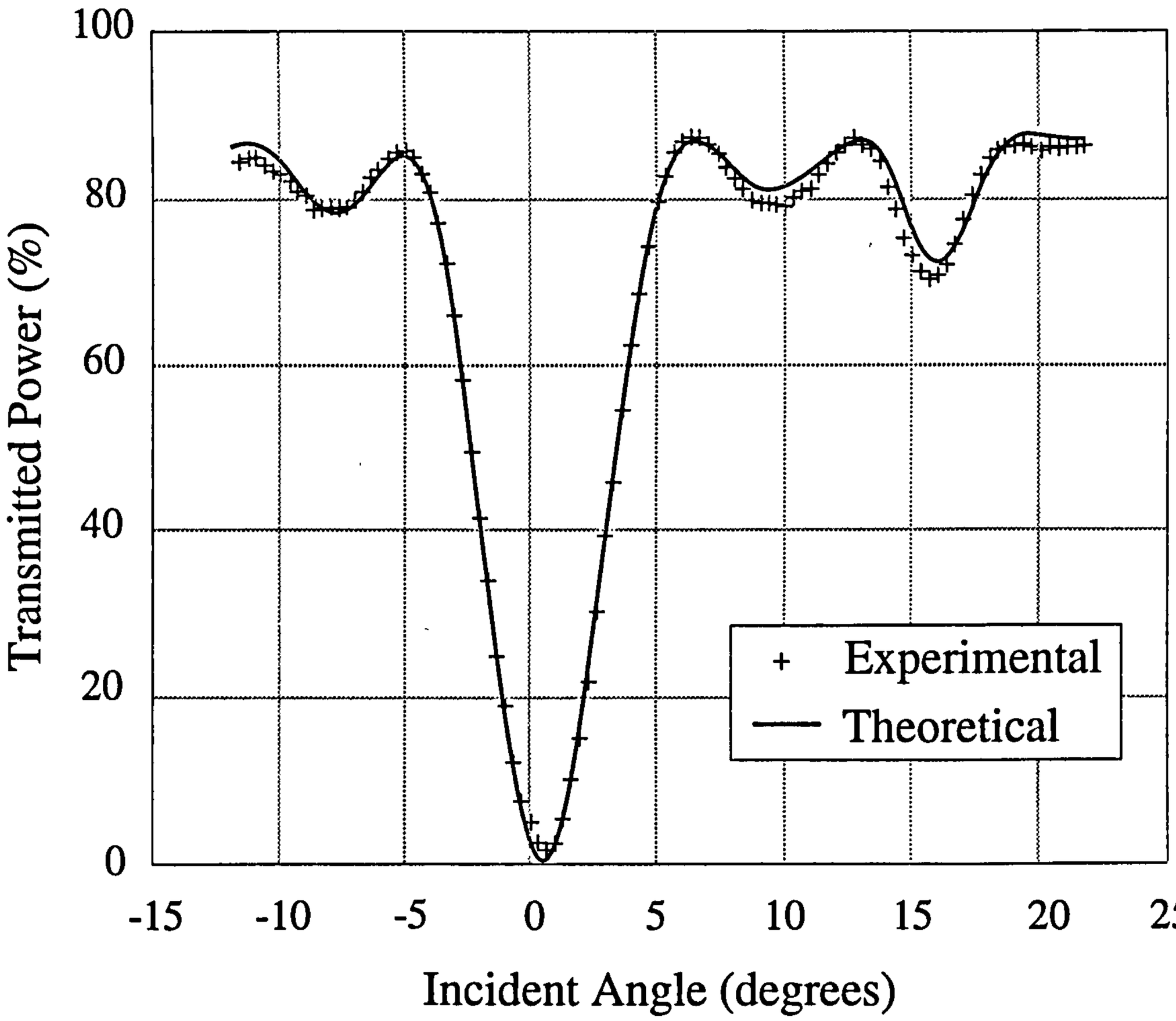


Figure 3.4: Multi-wave coupled-wave theory, of Redmond [RED89a], best fit to the experimental zero-order angular transmission spectrum of a volume transmission grating.

and uses a quadratic polynomial to model fringe plane and bulk refractive index non-uniformities. The processed gelatin had a thickness of 13μm, a refractive index of 1.493 and a scattering and/or absorption loss of 4.5%. The refractive index of the sensitised plate was measured to be 1.6 (for a 5% sensitiser concentration) and the amplitudes of the refractive index modulation of the fundamental and secondary grating harmonics were deduced (using the theory) to be 0.018 and 0.0045 respectively. Thickness changes in the grating were not directly incorporated into the results but a non-uniform grating vector was modelled. The results show an excellent correlation between experiment and multi-wave theory, clearly modelling the change in Bragg angle from 0° to 0.7°.

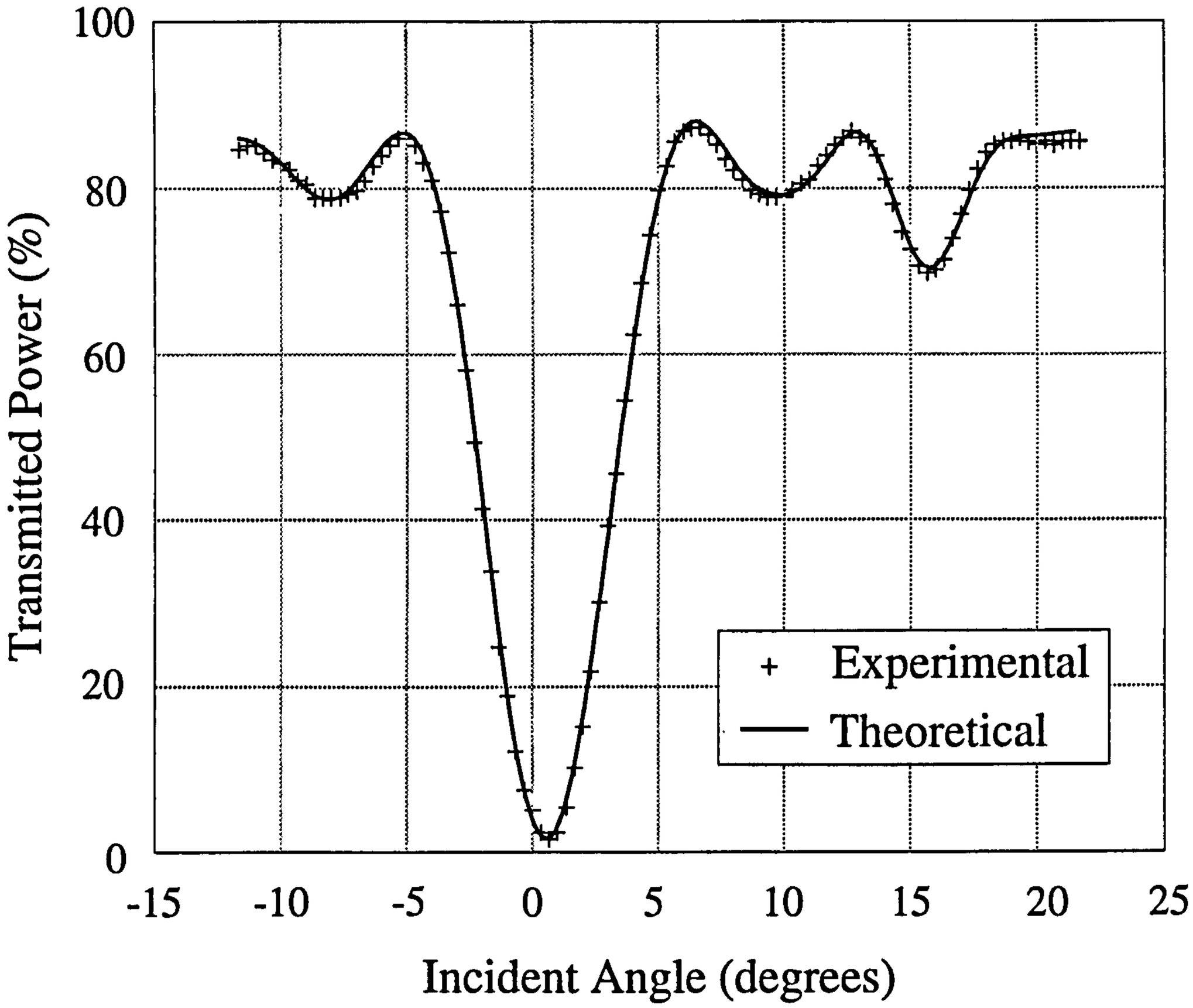


Figure 3.5: Rigorous coupled-wave theory (of this chapter) fit to the same volume transmission grating of Redmond [RED89a] shown in figure 3.4.

Rigorous model

Figure 3.5 shows the results of a RCWT fit to the same experimental data. The model incorporates the refractive indices and gelatin thickness measured by Redmond, with the inclusion of an absorption constant of $0.0034/\mu$m to account for the scattering and absorption losses. An emulsion shrinkage of 3% is required to correctly account for the Bragg angle change and a substrate is modelled as per section 3.6. The fit is achieved with refractive index modulation amplitudes of 0.0175 and 0.0071, for the fundamental and harmonic grating respectively, and maintains seven space-harmonics in the analysis.

Although seven diffracted orders were used in this analysis there was no appreciable difference in the numerical accuracy over the results using five space-harmonics.

This illustrates that the model has converged relatively quickly. However, the results retaining only three orders cannot account for the second harmonic grating, because the ± 2 diffraction orders are consequently not included in the model. The inclusion of nine or more orders introduces the numerical stability problems outlined in section 3.5.

The RCWT results are slightly better than those achieved using multi-wave theory and do not require the inclusion of bulk index and fringe structure non-uniformities. However, the physical parameters used to achieve this fit are slightly different to those measured by Redmond. Although still within experimental error they would suggest (particularly the thickness reduction) that small non-uniformities are present (in particular fringe angle variations with depth). In thicker gelatins (e.g. 40μm) the non-uniformities are more significant and need to be incorporated into the model when exact characterisation is required.

3.8.3 Reflection grating analysis

The terms reflection and transmission grating are to a certain extent misnomers. Depending upon the replay conditions, both reflected and transmitted diffraction orders are evident in the two types of grating. Optically generated holograms are more correctly referred to as predominantly reflective or transmissive subject to the recording conditions outlined in chapter 2. A pure transmission grating has a grating vector parallel to the surface ($\phi = 90°$), whereas a pure reflection grating has a grating vector orthogonal to the surface ($\phi = 0°$).

Moharam & Gaylord state quite explicitly in their original analysis that the RCWM will not work for pure reflection gratings (PRGs) in which the grating fringes are exactly parallel to the surface of the holographic medium, (such HOEs are often referred to as 'conformal' gratings [MOH81a, MOH81b]). If an exact model of such gratings is required, then they suggest a chain-matrix approach [MOH82a, RED89a]. However, Zylberberg & Marom [ZYL83] have come up with a mathematical formulation for PRGs that is identical to the RCWT, but in terms of the field expansion inside the grating it is not strictly a coupled-wave analysis. As a result of this the

algorithm requires many more diffracted orders to be included for numerical convergence [MOH83a]. This problem can, however, be viewed as being outweighed by the use of the same mathematical implementation; implying that the same analysis tool can be exploited to consider all gratings without the need for a separate theory. As stated by Moharam & Gaylord and corroborated by Zylberberg & Marom, a conformal grating can be adequately approximated in the RCWM by applying a very small angle to the grating (thus fulfilling the vector Floquet condition for field periodicity).

The grating chosen for analysis here is one that is used in chapter 6 as a nearest neighbour interconnect. It is pseudo-conformal, was designed to be recorded at 514.5nm and to operate at 1.064μm. The grating was recorded in a 40μm gelatin layer so that at the replay wavelength a normally incident plane wave is diffracted into the +1 reflected order, at an angle of 0.122° to the zero-order, (for exact details of the recording arrangement see section 6.2.3).

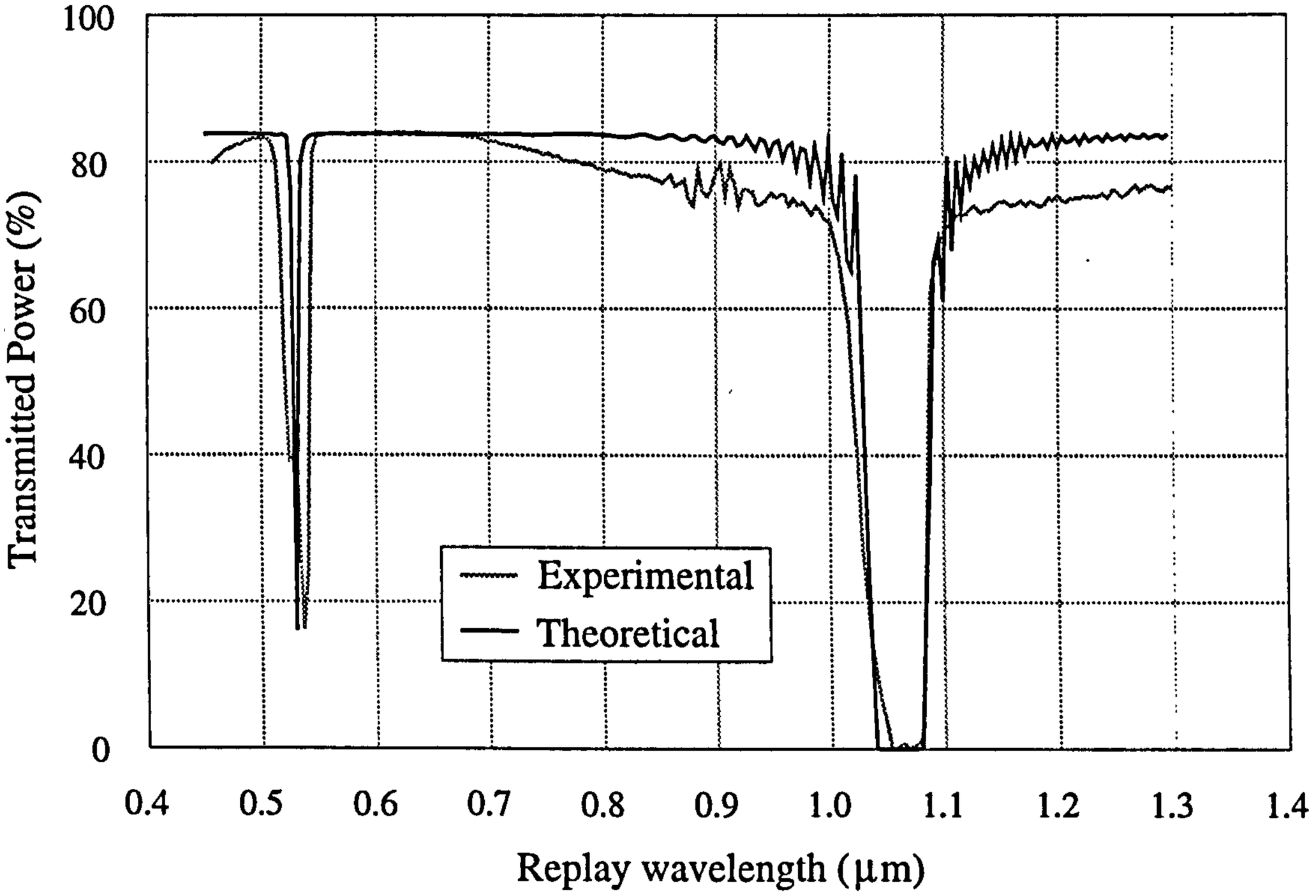


Figure 3.6: Rigorous coupled-wave theory (of this chapter) fit to the experimental transmitted wavelength spectrum of the zero-order of diffraction of a volume reflection grating.

Figure 3.6 shows the experimental and RCWM results for a wavelength scan of the zero-order transmitted power. The experimental results were obtained using a Shimadzu spectrophotometer and are not corrected for specular reflection or absorption. Newell [NEW87, pages 4-12 to 4-14] shows that bulk refractive index, index modulation and absorption in a processed hologram are all dependent upon the replay wavelength. This RCWM does not account for such variations which explains the discrepancy in the maximum zero-order power (in figure 3.6) with increasing wavelength. The grating was modelled using the substrate approximation, of section 3.6, and bulk refractive index and gelatin thickness changes were accounted for. Between the recording of this hologram and the writing of this chapter, the grating was no longer available to carry out exact thickness and refractive index measurements upon. However, the parameters used to model the grating: a change in index from 1.6 to 1.493, an absorption constant of 0.0013/μm and a gelatin swelling of 26%, are in agreement with the values measured by Newell for such gelatins (at 2.5% sensitiser concentration). This degree of swelling is wholly expected since the grating is almost a PRG and as such will sustain a significantly greater degree of swelling than a slanted grating, (due to the lack of mechanical rigidity inherent in a PRG) [NEW87, page 3-18]. The spectrum shows a strong secondary minimum at about half the designed replay wavelength. From this data the model estimates fundamental and secondary index modulation amplitudes of 0.06 and 0.0083 respectively. These were arrived at with the experimental evidence of the wavelength scan and an experimental determination of the angular bandwidth (discussed in section 6.2.4).

The model included twenty-one space-harmonics which illustrates that a large number of harmonics can be supported by the model in certain circumstances. However, the numerical accuracy was no greater than the results obtained using a mere five orders which again illustrates the speed of convergence of the algorithm. It can be seen that the results for this relatively thick gelatin hologram are not as accurate as the transmission case presented in section 3.8.2. This is due in part to the presence of grating and bulk index non-uniformities (which tend to asymmetrically broaden the bandwidth at the harmonic wavelengths) and to the possible existence

of noise gratings discussed by Keininen & Salminen [KEI93]. Although there are minor discrepancies in the theoretical fit, the theory accurately predicts all the necessary physical parameters required when designing a reflection HOE interconnect, including angular and wavelength bandwidths.

3.9 Summary

The rigorous coupled-wave method is a very flexible and accurate tool. It shows very good agreement with experiment for both transmission and reflection planar gratings. The numerical instabilities that can be evident [CHA94] have been shown not to be too restrictive. Although exact characterisation of all types and thickness of DCG grating are not achievable it is not the aim of this model to do this. It is more than adequate in predicting and analysing the key physical characteristics of transmission and reflection optical interconnects fabricated in the form of DCG holograms.

The comparison drawn between the multi-wave theory of Redmond [RED89a] and this RCWM for relatively thin gelatins suggests that the non-uniformities in bulk refractive index and grating vector may not be as pronounced as previously estimated. To exactly characterise thick and thin gelatin holograms, a theory in which the grating vector can be variant in the $z-$direction (to allow for non-uniformity polynomials) and which is wholly rigorous would be required. BKK being a multilayer rigorous model would appear to be the best candidate if such a characterisation were to be carried out.

Chapter 4

Wavelength Shifting

DCG HOEs, whether they are transmission or reflection, can be split into two broad categories: planar gratings and those with focusing power. Fabricating either of these two types of grating for replay at the recording wavelength is straightforward. No aberrations or change in replay conditions will occur if the physical properties do not alter during processing.

Due to the limited spectral sensitivity of DCG, it is not possible to directly record holograms at wavelengths greater than 530nm. However, most of the optical communications and computational systems which have been researched to date operate in the near infra-red (IR). Planar transmission and reflection gratings can be made with relative ease at shorter wavelengths (typically 488 and 514 nm) for operation in the red and near IR. If, however, gratings with focusing power are required, aberrations are induced by the wavelength shift as a result of diffractive dispersion. Even comparatively small wavelength shifts can produce difficulties when large apertures are needed, such as the important application of aircraft head-up displays (HUDs) [CLO75, WOO83, EVA89]. HUDs must work with the phosphor emission line of a cathode ray tube (543nm), and although this is only just above the sensitivity threshold of DCG, the required wavelength-shifted recording geometry must be carefully optimised, (usually requiring aspheric wavefronts).

When designing wavelength-shifted DCG gratings the problem may be conveniently divided into: geometrical diffraction and volume (Bragg) effects. Although

these two are linked they can be considered separately. Geometrical diffraction affects concern the way an HOE redirects and shapes the incident wavefront whereas volume effects mainly alter how efficiently the wavefront is transformed [CLO75, page 409].

For planar gratings the geometrical considerations are limited simply to the incident and diffracted angles of a plane wavefront. The Bragg angle match will be the same across the hologram and so the volume considerations are equally straightforward.

Gratings with focusing power introduce geometrical aberrations. As the fringe period varies across the hologram surface, the Bragg condition can no longer be satisfied at all points on the hologram at the same time. This means that the efficiency of the hologram varies from point to point. Consequently, when designing wavelength-shifted gratings with focusing power, there is generally a trade-off between the geometrical and volume considerations.

This chapter examines both the geometric and volume considerations of planar and focusing gratings in both transmission and reflection:

- **Section 4.1** analyses the geometric diffraction considerations of planar gratings.
- **Section 4.2** examines the nature of the dispersive aberration in gratings with focusing power and outlines the basic approaches to overcoming this problem, (including the efficiency trade-offs).
- **Section 4.3** presents a new and simple method for designing and recording wavelength-shifted gratings with focusing power.
- **Section 4.4** assesses the usefulness of the new method when applied to wavelength-shifted transmission HOEs.
- **Section 4.5** tackles the volume efficiency considerations, outlining a method of achieving maximum efficiency in either type of grating in reflection or transmission; and

- **Section 4.6** describes some results of recording wavelength-shifted spherical and cylindrical focusing elements in DCG and presents a method of making such elements operate in-line and on-axis.

4.1 Planar gratings

In this chapter the internal angle convention (see section 1.5) is dispensed with and the external angle convention is adhered to throughout. This is done because it is more usual and convenient to describe recording and replay angles, θ, for HOEs in a range $-90° \leq \theta \leq 90°$. Although this leads to slightly different equations for transmission and reflection gratings it is the external recording angles that need to be solved for in a wavelength-shift (see section 4.1.1) making the external angle convention more appropriate.

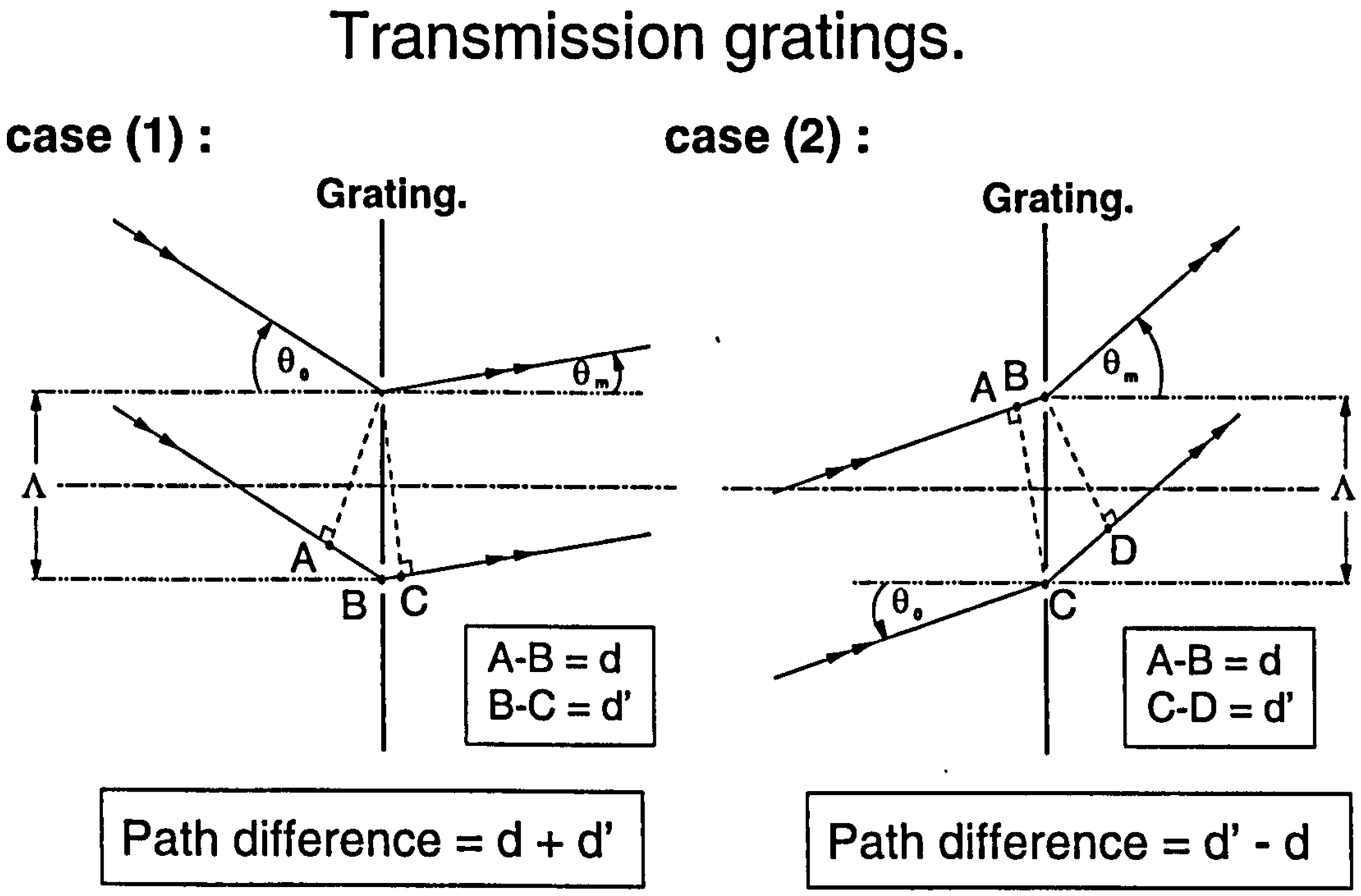


Figure 4.1: Ray paths for the two possible transmission grating geometries.

Consider either of the transmission gratings shown in figure 4.1. The optical path difference between the plane wave incident upon the grating and the resultant

diffracted plane wave is given by:

$$\begin{aligned} \text{path difference}\,\{\text{case(1)}\} &= d + d' \text{ ; and} \\ \text{path difference}\,\{\text{case(2)}\} &= d' - d \end{aligned} \tag{4.1}$$

The incident and diffracted waves must be in phase for a diffraction maximum, so that in case (1)

$$d + d' = m\lambda = \Lambda \sin|\theta_m| + \Lambda \sin|\theta_0| \,, \tag{4.2}$$

where λ is the wavelength, Λ is the grating period, m is the diffraction order, θ_0 is the angle of the incident plane wave and θ_m is the angle of the subsequent diffracted wave. In case (2)

$$d' - d = m\lambda = \Lambda \sin|\theta_m| - \Lambda \sin|\theta_0| \tag{4.3}$$

Applying the external angle sign conventions (see section 1.5) and the fact that we are concerned with the $m = -1$ diffracted order (as per chapter 3) these two equations reduce to the single transmission grating diffraction equation:

$$\sin\theta_0 - \sin\theta_{-1} = \frac{\lambda}{\Lambda} \tag{4.4}$$

Similarly for the two cases of a planar reflection grating, shown in figure 4.2, we get the reflection grating diffraction equation:

$$\sin\theta_0 + \sin\theta_{-1} = \frac{\lambda}{\Lambda} \tag{4.5}$$

When Bragg effects are considered the incident ray angle relative to the fringe plane becomes important. To meet the Bragg condition, incident and out-going angles must be equal, (symmetric about the grating vector, **K**). Consider the transmission grating Bragg planes shown in figure 4.3. Here again the two incident rays must have optical paths that add in phase:

$$\text{path difference} = d + d' = m\frac{\lambda}{n} \,, \tag{4.6}$$

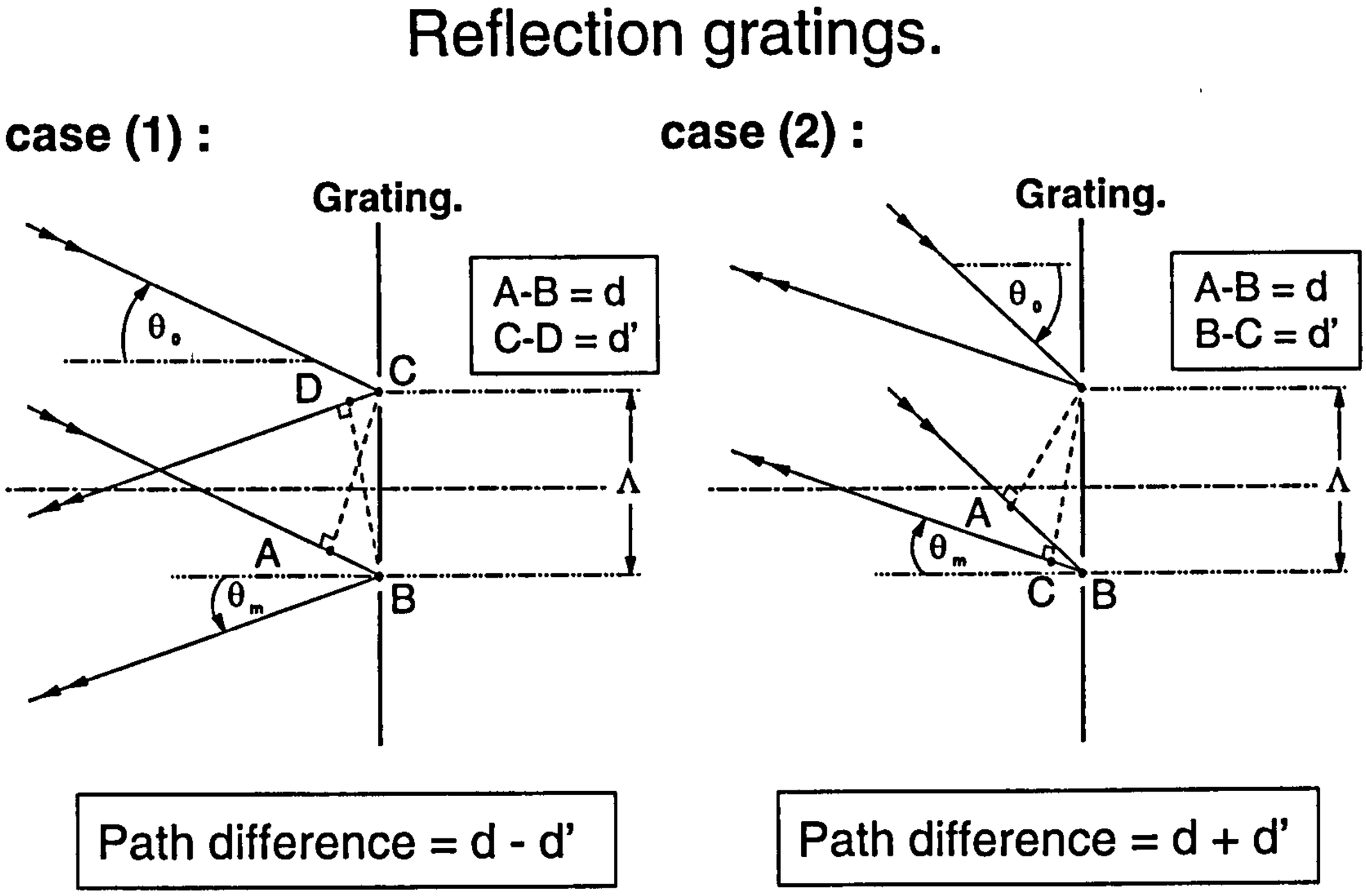


Figure 4.2: Ray paths for the two possible reflection grating geometries.

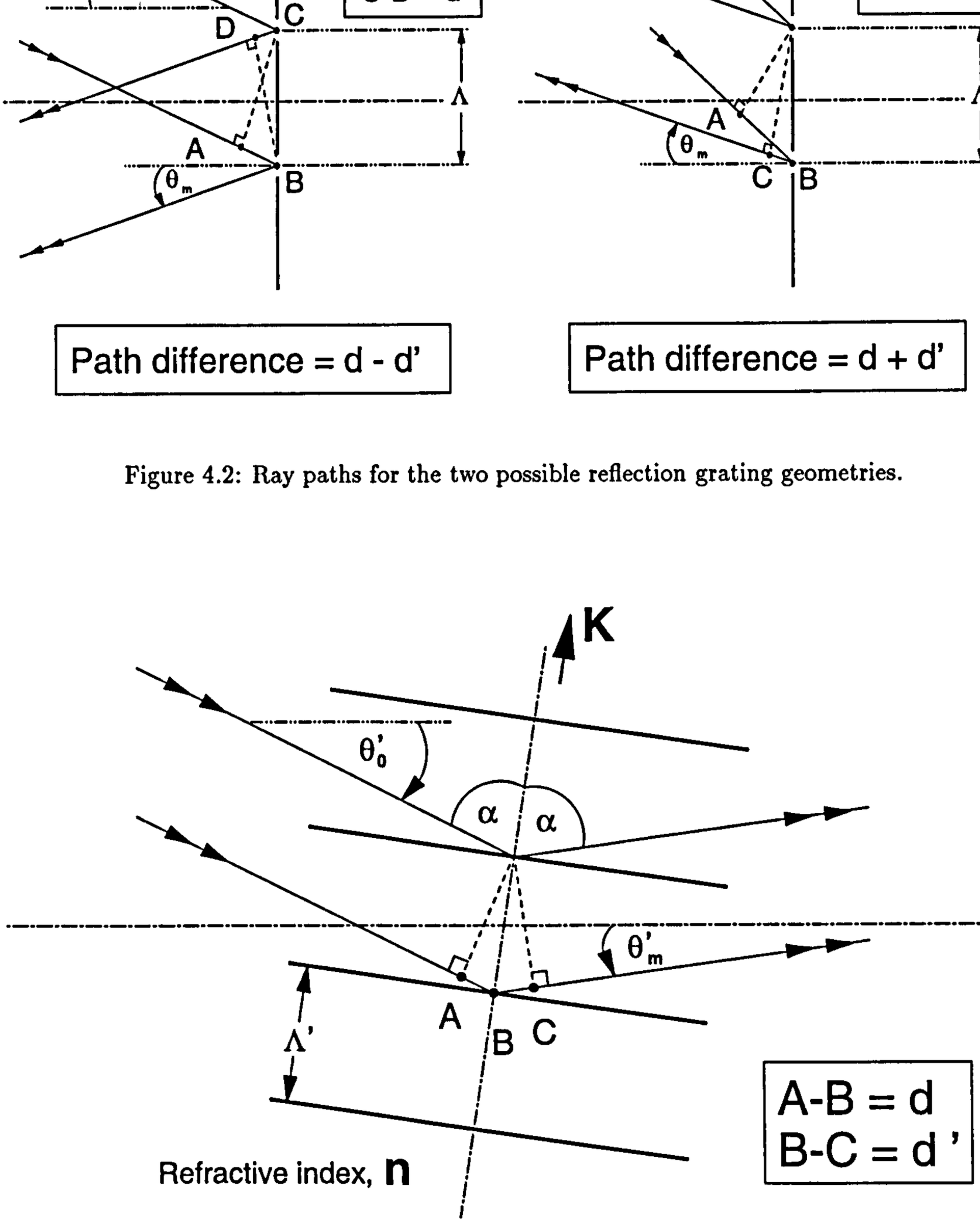


Figure 4.3: Diffracted ray paths from the Bragg planes of a transmission grating.

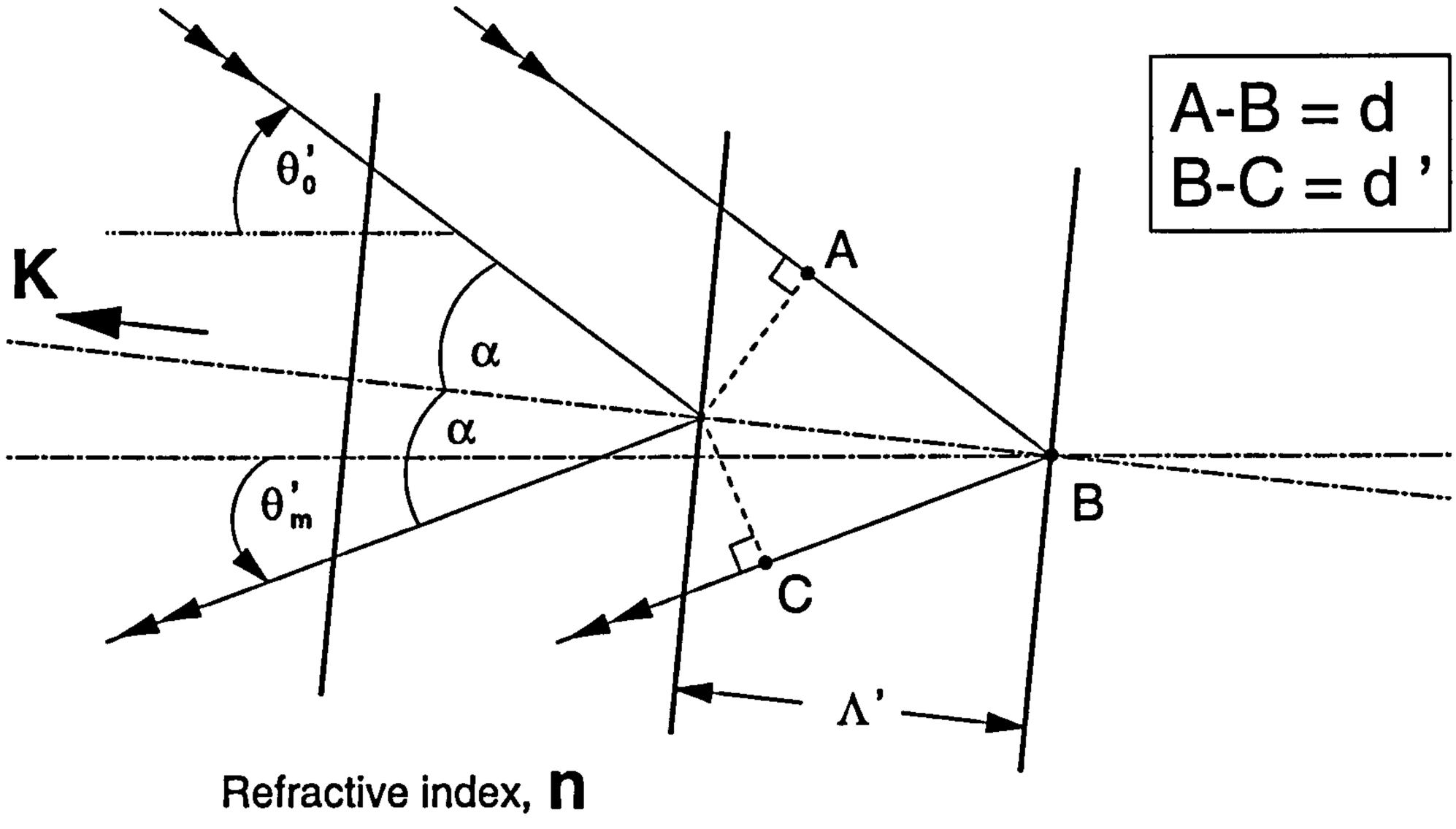


Figure 4.4: Diffracted ray paths from the Bragg planes of a reflection grating.

where n is the refractive index of the volume holographic material. When expressed with respect to the Bragg planes this becomes

$$2\Lambda' \cos\alpha = m\frac{\lambda}{n} \tag{4.7}$$

Reformulating equation 4.7 in terms of the incident and diffracted angles and adhering to sign conventions, it becomes the Bragg equation for planar transmission gratings:

$$\sin\left(\frac{\theta'_{-1} - \theta'_0}{2}\right) = \frac{\lambda}{2n\Lambda'} \tag{4.8}$$

Similarly, for the reflection grating Bragg planes shown in figure 4.4, the Bragg equation for planar reflection gratings is:

$$\cos\left(\frac{\theta'_{-1} - \theta'_0}{2}\right) = \frac{\lambda}{2n\Lambda'} \tag{4.9}$$

The transmission grating equation 4.4 describes the input and diffracted wave angles and the transmission Bragg equation 4.8 describes the Bragg condition ge-

ometry. When the Bragg condition is met, the strongest coupling between incident and diffracted waves occurs. Similarly equations 4.5 and 4.9 define the reflection diffraction geometries. These two sets of equations form the basis for recording wavelength-shifted planar gratings and, as described in section 4.3, can form a basis for gratings with focusing power.

4.1.1 Wavelength-shifted planar gratings

The four equations 4.4, 4.8, 4.5 and 4.9 can be further subdivided into two sets of equations for the replay and recording geometries. To distinguish between recording and replay variables a notation is adopted such that any variable that refers to recording has a zero suffix on the subscript and those that refer to replay have a suffix of one; e.g. θ_{10} is a recording variable θ_1; whereas θ_{01} is a replay variable θ_0.

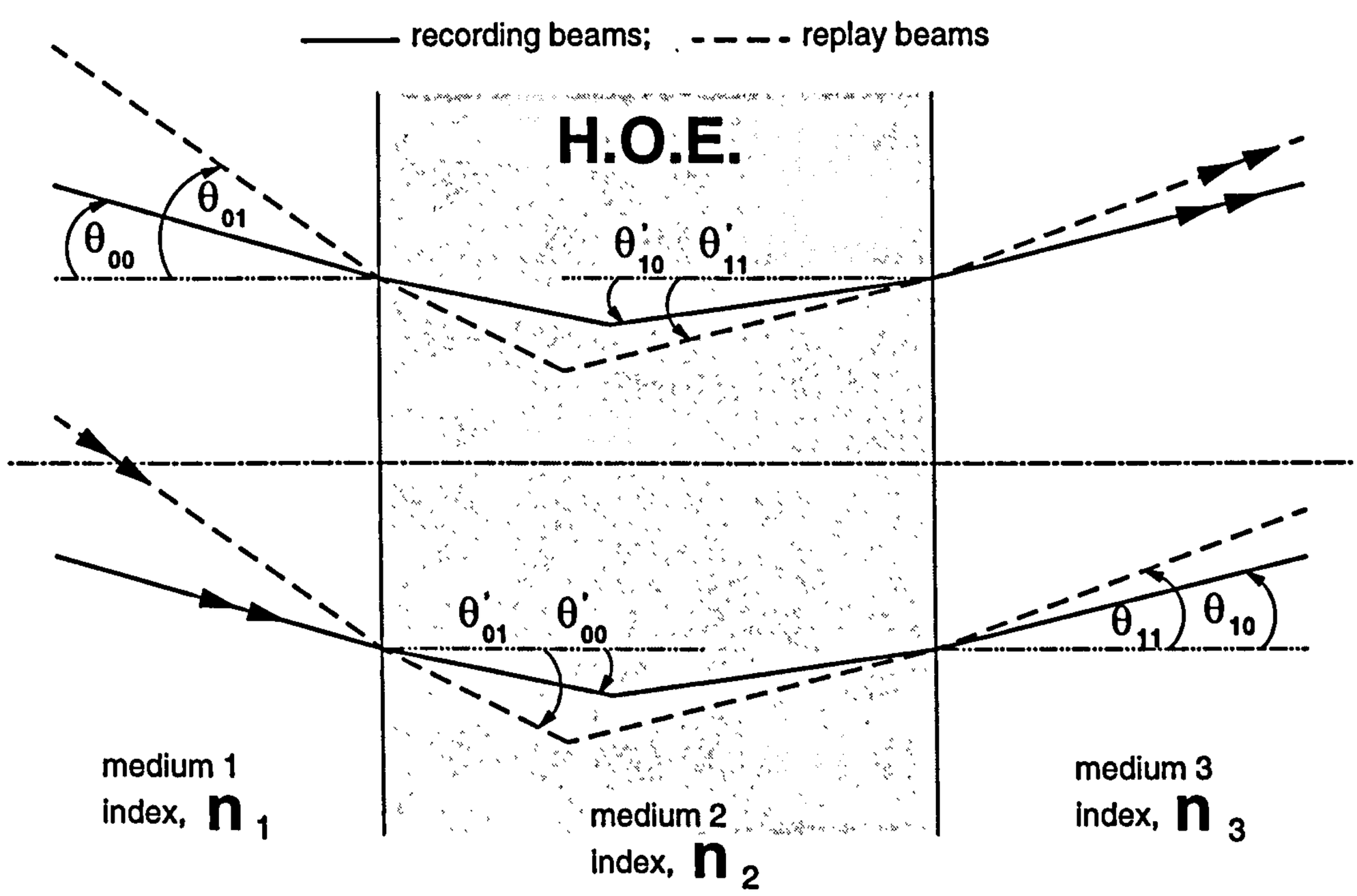


Figure 4.5: Recording and replay geometries for a wavelength-shifted transmission HOE.

Transmission gratings

The transmission grating diffraction equations for recording and replay can be expressed in terms of the surface fringe period, Λ, and then equated with one another, since this period is fixed between recording and replay geometries in the holographic recording material. Figure 4.5 shows the respective recording and replay geometries for a transmission grating. Considering this with the diffraction equation 4.4 gives

$$\frac{n_{20}}{\lambda_0}\left[\sin\theta'_{00} - \frac{n_{30}}{n_{10}}\sin\theta'_{10}\right] = K_A = \frac{n_{21}}{\lambda_1}\left[\sin\theta'_{01} - \sin\theta'_{11}\right] , \qquad (4.10)$$

where $K_A = {}^1/\Lambda$, n_1 is the refractive index of medium 1, n_2 is the refractive index of the hologram, n_3 is the refractive index of medium 3 and all angles have the same meaning as in section 4.1, (the suffixes to these variables again denote recording or replay). Similarly, the transmission grating Bragg equations for recording and replay can be expressed in terms of the internal fringe period, Λ', and then equated with one another. The Bragg equation 4.8 examined with respect to figure 4.5 gives

$$\frac{n_{20}}{\lambda_0}\left[\sin\left(\frac{\theta'_{10} - \theta'_{00}}{2}\right)\right] = K_B = \frac{n_{21}}{\lambda_1}\left[\sin\left(\frac{\theta'_{11} - \theta'_{01}}{2}\right)\right] , \qquad (4.11)$$

where $K_B = {}^1/2\Lambda'$. For a particular replay condition the two parameters of interest in equations 4.10 and 4.11 are the two recording beam angles, θ'_{00} and θ'_{10}, these can be solved for as follows:

$$\theta'_{10} = 2\sin^{-1}\left(\frac{\lambda_0 K_B}{n_{20}}\right) + \theta'_{00} \qquad (4.12)$$

and

$$\sin\theta'_{00} = \frac{-b \pm \sqrt{(b^2 - 4ac)}}{2a} , \qquad (4.13)$$

where

$$\begin{aligned} a &= \chi_A^2 + \chi_B^2 , \\ b &= -\frac{2\lambda_0 \chi_A K_A}{n_{20}} , \end{aligned}$$

$$
\begin{aligned}
c &= \frac{\lambda_0^2 K_A^2}{n_{20}^2} - \chi_B^2 \,, \\
\chi_A &= 1 + \frac{n_{30}}{n_{10}} \cos\left(2\cos^{-1}\left(\frac{\lambda_0 K_B}{n_{20}}\right)\right) \text{ ; and} \\
\chi_B &= \frac{n_{30}}{n_{10}} \sin\left(2\sin^{-1}\left(\frac{\lambda_0 K_B}{n_{20}}\right)\right)
\end{aligned} \tag{4.14}
$$

The quadratic nature of the solution to equations 4.10 and 4.11 leads to two possible values for θ'_{00}, (from equation 4.13). Only one of these values is of use, the other either makes no physical sense or is the value of θ'_{10}. The exterior recording angles θ_{00} & θ_{10} are obtained using Snell's law.

Reflection gratings

The reflection grating diffraction equations can be equated in a similar way:

$$
\frac{n_{20}}{\lambda_0}\left[\sin\theta'_{00} + \frac{n_{10}}{n_{30}}\sin\theta'_{10}\right] = K_A = \frac{n_{21}}{\lambda_1}\left[\sin\theta'_{01} + \sin\theta'_{11}\right] \tag{4.15}
$$

The reflection Bragg equations give

$$
\frac{n_{20}}{\lambda_0}\left[\cos\left(\frac{\theta'_{10} - \theta'_{00}}{2}\right)\right] = K_B = \frac{n_{21}}{\lambda_1}\left[\cos\left(\frac{\theta'_{11} - \theta'_{01}}{2}\right)\right] \tag{4.16}
$$

The recording beam angles, θ'_{00} and θ'_{10}, can be solved for using equations 4.15 and 4.16 above to give a similar result to transmission gratings:

$$
\theta'_{10} = 2\cos^{-1}\left(\frac{\lambda_0 K_B}{n_{20}}\right) + \theta'_{00} \tag{4.17}
$$

and

$$
\sin\theta'_{00} = \frac{-b \pm \sqrt{(b^2 - 4ac)}}{2a} \,, \tag{4.18}
$$

where

$$
\begin{aligned}
a &= \chi_A^2 + \chi_B^2 \,, \\
b &= -\frac{2\lambda_0 \chi_A K_A}{n_{20}} \,,
\end{aligned}
$$

$$
\begin{aligned}
c &= \frac{\lambda_0^2 K_A^2}{n_{20}^2} - \chi_B^2 \,, \\
\chi_A &= 1 + \frac{n_{10}}{n_{30}} \cos\left(2\cos^{-1}\left(\frac{\lambda_0 K_B}{n_{20}}\right)\right) \;; \text{ and} \\
\chi_B &= \frac{n_{10}}{n_{30}} \sin\left(2\sin^{-1}\left(\frac{\lambda_0 K_B}{n_{20}}\right)\right)
\end{aligned} \tag{4.19}
$$

Equations 4.12, 4.13 and 4.14 for transmission gratings and the counterpart equations 4.17, 4.18 and 4.19 for reflection gratings represent a solution to wavelength shifting planar gratings which is easily coded into a computer program.

4.2 Gratings with focusing power

The limited spectral sensitivity of DCG makes it impossible to directly record holograms with focusing power at red and near IR wavelengths. Recording a grating with focusing power at a different wavelength than that at which it is intended to replay induces aberrations as a result of diffractive dispersion. This type of chromatic aberration is unique to volume HOEs. An understanding of the problem can be seen with reference to figure 4.6.

A volume holographic lens of thickness D recorded at a wavelength λ_g between a plane wave at an angle θ_g and a spherical wave converging on a focal point at f_g, when replayed with the same plane wave at the same wavelength will diffract all rays to the same focal point f_g. This volume element can be considered as a large number of cascaded thin zone lenses, each with a slightly different focal length, (see [HEC87, pages 445 and 595] and [STO86, pages 32-39]). For the front face and rear face zone lenses the focal lengths can be expressed as:

$$f_{\lambda_g}^{front} = f_g \tag{4.20}$$

and

$$f_{\lambda_g}^{rear} = f_g - D \tag{4.21}$$

respectively. Such thin holographic lenses (indeed all holographic gratings with focusing power) exhibit focal lengths that are inversely proportional to the wavelength

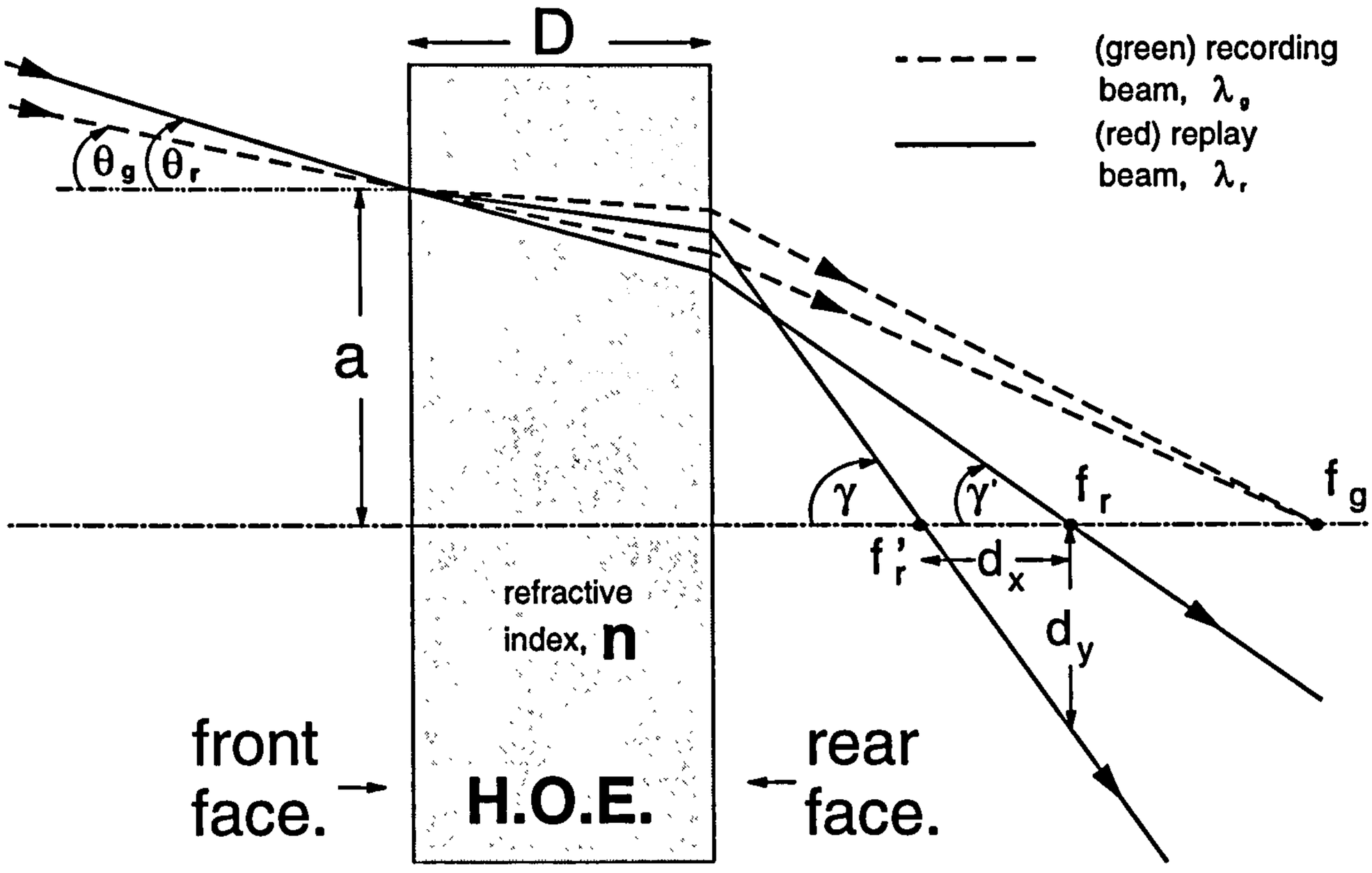


Figure 4.6: Dispersion problems with wavelength-shifted focusing HOEs.

at which they are used, [HEC87, page 446]. Thus at a replay wavelength λ_r, where a is small in comparison to f_g, these focal lengths become can be approximated to:

$$\begin{aligned} f_r &= f_{\lambda_r}^{front} \approx \frac{f_g \lambda_g}{\lambda_r} \text{ ; and} \\ f_r' &= f_{\lambda_r}^{rear} \approx \frac{(f_g - D)\lambda_g}{\lambda_r} \end{aligned} \tag{4.22}$$

Exact expressions for f_r and f_r' are given by the grating equation, but for the qualitative basis of this discussion the approximations will suffice. Clearly the front and rear face grating profiles of this volume HOE focus rays into different focal points when replayed at a wavelength different to that at which it was recorded. This creates a type of chromatic aberration in the focal point which has both transverse, d_y, and longitudinal, d_x, components, where

$$d_x = f_r - f_r' \approx D\frac{\lambda_g}{\lambda_r} \tag{4.23}$$

and

$$d_y = d_x \tan\gamma \approx D\frac{\lambda_g}{\lambda_r}\left(\frac{a - D\tan\left(\sin^{-1}\left(\frac{\sin\theta_r}{n}\right)\right)}{f_r - D}\right) \tag{4.24}$$

By examining equation 4.24 it is clear that the magnitude of the transverse dispersive aberration depends upon:

- the wavelength shift,
- the thickness of the volume holographic medium,
- the average Bragg diffraction angle; and
- the numerical aperture of the holographic lens.

Note that these aberrations scale with the physical size of the system, whereas the diffraction limit does not, so the relative importance of these geometrical effects varies with HOE size [RED89a, page 185].

4.2.1 Methods of solution

Many papers were published in the late 60s and early 70s on the subject of quantifying aberrations in holographic lenses, [LEI65, CHA67, LAT71a, LAT71b, LUD73, WEL75]; Close [CLO75] presents a detailed overview. Some of these are analytical and others involve analytical ray tracing, but none of them present techniques for minimising the aberrations.

Methods for minimising volume holographic lens aberrations on reconstruction with a wavelength shift vary considerably in their approach, largely due to the trade-off which exists between aberration minimisation and efficiency optimisation.

Each method achieves a different degree of 'Bragg matching', either optimising wholly for aberrations and obtaining poor Bragg matching, obtaining a good Bragg match but poor aberration control or some halfway house between these two extremes. Although the methods are diverse they all adopt one or more of the following approaches:

1. Apply to a generalised point to point off-axis replay geometry *OR* place more

specific constraints upon the replay conditions; such as an off-axis plane wave coming to a focus on-axis.

2. Adopt an analytical (geometrical optical) *OR* an empirical ray tracing approach.

3. Optimise for Bragg matching *OR* aberration minimisation.

4. Involve a simple spherical plane wave interferometric recording procedure *OR* require the generation of aspheric wavefronts in the recording process; usually implying that high space bandwidth product (SBWP) CGHs must be designed and fabricated first.

5. Involve a single stage recording process *OR* a recursive recording technique; and

6. Consist of a single volume hologram *OR* cascades of two or more holograms.

The following represent a broad spectrum of the optimisation and recording techniques available:

- **Lin and Doherty** [LIN71] describe an approach which is applicable to a generalised replay geometry. It involves a two stage recording process by firstly recording the desired HOE at the required replay wavelength in a medium which is sensitive at that wavelength, (such as dye sensitised silver halide emulsion). The HOE is then contact copied into DCG in the blue or green (488 or 514 nm) in an attempt to get a higher diffraction efficiency. This results in a single volume HOE which has exactly the right surface fringe profile and therefore excellent aberration performance. By careful analytical optimisation of the copying geometry and DCG thickness a high diffraction efficiency can be achieved. The process does not require CGHs and can be accomplished with simple spherical and planar wavefronts. The aberrations are optimised and Bragg matching is approximated under the constraints of the copying process. This technique is limited to situations where a laser and suitable (first stage) recording material are available at the replay wavelength. In addition to this,

the silver halide emulsion can be prone to a high degree of scatter and have very low sensitivity at IR wavelengths.

- **Latta and Pole** [LAT79] describe an approach which is applicable to a generalised replay geometry. It involves a single stage recording process using simple spherical and planar wavefronts. The method optimises the Bragg matching of the HOE using **K**-vector closure so that the average Bragg condition is fulfilled. It is a ray tracing approach whereby the hologram is divided into small uniform gratings meeting the average Bragg condition, (on replay). Rays corresponding to Bragg incidence at recording are calculated for each sub-division and traced for the entire hologram to find the minimum blur circle points. The two resulting points correspond to the best position for the recording point sources. This results in a single hologram which is efficient but does not optimise the aberration performance.

- **Winick** [WIN82] describes a generalised approach that involves a single step process using CGHs to generate the aspheric wavefronts needed during recording. This analytical method calculates the exact object and reference beam phases at the recording wavelength needed to produce diffraction limited aberration performance at the replay wavelength. CGHs are fabricated to produce these wavefronts. The Bragg condition is approximated by choosing a recording geometry in which the fringe slant inside the HOE is as close as possible to that required for 100% diffraction efficiency without breaking the constraints of aberration phase matching. This results in a single hologram with diffraction limited performance and high efficiency. The CGHs required, however, generally require a high SBWP and as such are difficult to fabricate. CGHs themselves can offer a flexible solution to designing and producing optimised HOEs with focusing power at a specific wavelength. Recent advances in multi-level high efficiency CGHs ([MIL93, chapter3] and [NOP93]) make this a realistic alternative to volume holography, (see [LEE87] for a review of CGH techniques).

- **Stone and George** [STO85] describe a method which is limited to minimis-

ing longitudinal dispersive aberrations (see equation 4.23), with the aim of producing achromatic holographic lens systems with a wide wavelength and angular response. A ray tracing approach is adopted to optimise a set of holographic lens focal lengths and separation distances in order to realise this aim. In addition, the model includes hologram efficiency calculations to maximise the wavelength and angular bandwidth. This results in a set of two or more cascaded holograms which can be straightforwardly recorded with spherical and plane wavefronts. Each hologram in the cascade must be tuned to match the others in their Bragg response so as to achieve the maximum efficiency and bandwidth criteria.

- **Herzig** [HER86] describes a two stage recording method for producing off-axis holographic lenses with a large f/number. These were for a specific use with semiconductor lasers and were designed to remove their inherent astigmatism. The approach is analytical, involving the recording of astigmatic intermediate holograms which are used to generate the wavefronts required to record the final HOE in a geometry designed to optimise Bragg matching. This results in a single high efficiency hologram, does not require CGH recording and produces non-astigmatic wavefronts when used with the diode laser it was designed to compensate. **Prongue and Herzig** [PRO89] generalise and simplify this approach by analytically calculating recording points based on minimising the 2nd and 3rd order geometric aberrations. An intermediate HOE is recorded using spherical wavefronts at the calculated recording points and this is contact copied onto the final hologram. Bragg matching is achieved by optimising the copying geometry in a similar way to Lin and Doherty. The resultant hologram is both efficient and diffraction limited, but only for small numerical apertures.

- **Chen, Hershey and Leith** [CHE87] describe an analytical procedure suitable for designing and recording small f/number on-axis holographic lenses by balancing the 3rd against the 5th order geometrical aberrations. A CGH phase-compensator plate is manufactured and placed in one arm of the interferometric recording arrangement with an appropriate spherical wavefront in

the other to achieve the desired surface fringe profile. The recording geometry is optimised for Bragg matching and yields a single high efficiency aberration-minimised holographic lens. The compensator plate can be recorded holographically, removing the need for a high SBWP CGH. Apertures as large as 5cm at $f/1.8$ have been demonstrated for 633nm making this a very versatile solution to the wavelength shifting problem.

- **Assenheimer, Amitai and Friesem** [ASS88] describe a method for focusing a collimated off-axis beam to an on-axis point using a large aperture HOE. The technique is analytical and involves removing all Seidel (3rd order geometrical) aberrations by generating aspheric wavefronts with which to record the hologram. It is novel in that it involves a recursive technique for generating these aspheric wavefronts from ancestor holograms recorded with spherical wavefronts. The final holographic lens is recorded in a Bragg optimised geometry in order to maximise efficiency on replay.

- **Redmond** [RED89a] describes a purely ray traced solution which is general in its applicability. The recording process is single step using spherical wavefronts and produces a highly efficient single element hologram. The method sub-divides the HOE into smaller sectors each assumed to be a uniform planar volume grating. The grating vector for each sub-division is calculated based upon the recording points currently being considered in the optimisation, (these can initially be random). Rays are traced through each sector according to the replay conditions and the calculated grating vectors. The rays exiting the HOE are traced through a specified distance while the spot dimension is regularly calculated, (effectively a focal plane scan). In this way the best focal spot position is found. The recording geometry is optimised by a simple nested search of possible recording regions. Once the best region is identified, the search switches to that region which is divided up and searched, and so on. This 4-dimensional search is a brute force solution, but quite suitable for modern desktop computers. The method is versatile in that the Bragg mismatch for each hologram sector can also be calculated and minimised, al-

lowing a trade-off between optimum power and optimum power density.

Most of these methods involve fulfilling the average Bragg condition for the wavelength shift and minimising aberrations in some way. Those utilising CGHs to produce the exact aspheric wavefronts are the most general and effective. All of these techniques, however, are complicated by multi-stage recording processes or CGH production. If arrays of lenslets with different replay geometries per lenslet are required, (as with a complex optical interconnect - see chapter 7), a simpler design and recording technique is required.

A simple alternative method is presented in the following section and its general applicability analysed. For the range of interconnects used in this work the method is found to produce lenslets which operate close to or within the diffraction limit. It also produces recording geometries in agreement with those of Redmond (within experimental error).

4.3 Recording geometry calculation method

Appendix B presents a mathematical analysis of the form that holographic fringes take through the depth of the medium and on the surface, for a HOE with focusing power. The conclusions are similar to those of Stone [STO86], for on-axis diffractive lenses, showing that the surface structure is elliptical and closely resembles a thin Fresnel zone plate (FZP).

The polynomial B.5 used to determine fringe structure could in theory be used to generate the data necessary to record wavelength-shifted lenslets. The algebraic form, however, is cumbersome for this purpose and a more convenient way of looking at the problem is as follows. The surface holographic fringes are elliptical and resemble a FZP. Indeed, a zone plate type of analysis can be applied to them. Consider a thin (medium thickness $\ll f_0$ or f_1) elliptical fringe, diffractive structure in cross section (ignoring volume effects for the time being), such a structure will look like figure 4.7. The co-ordinate system used here is that specified in section 1.5, however, the focal distances f_0 and f_1 will always be positive and a distinction needs to be made in order to specify whether points R and S lie to the left or the right of

the HOE. Consequently the convention adopted is the same as that used in geometric optics for thin lenses; that is, f_0 is positive to the left of the HOE and negative to the right, (the opposite applies to f_1).

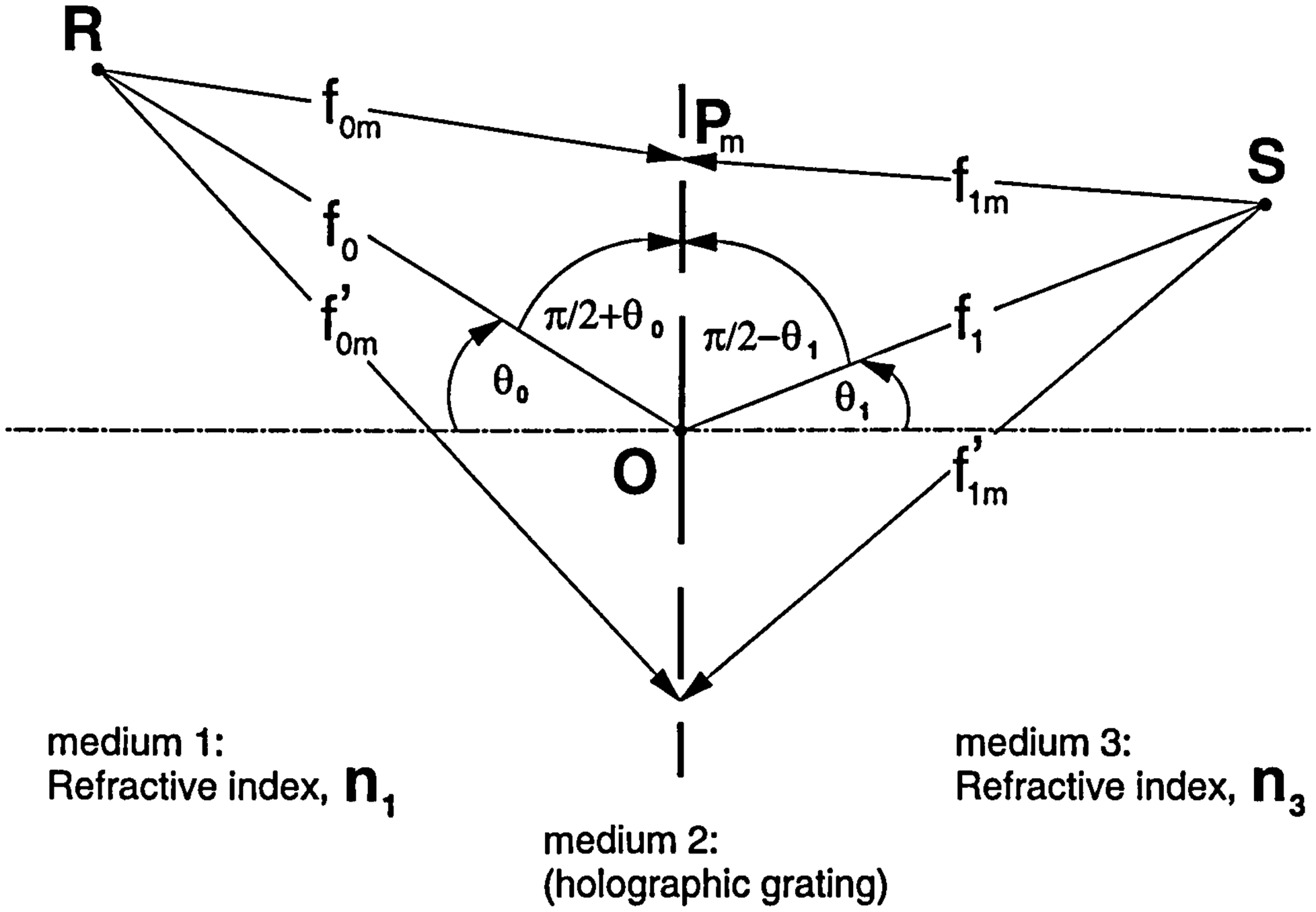


Figure 4.7: Off-axis zone plate style geometry for two point sources, (note that the external angle convention (section 1.5) is applied, hence the angle $\pi/2 + \theta_0$).

Consider the two optical paths of rays from points R and S: when replaying the grating structure depicted (at the same wavelength at which it was recorded), a ray traversing the path $R - O - S$ must, by definition, be an integer number of half wavelengths out of phase with a ray following $R - P_m - S$, (where point P_m marks the outer edge of the m:th zone). Hence

$$(n_1 f_{0m} + n_3 f_{1m}) - (n_1 f_0 + n_3 f_1) = \frac{m\lambda}{2}\,, \tag{4.25}$$

where m is an integer corresponding to the m:th zone, λ is the wavelength, n_1 and n_3 are the refractive indices of mediums 1 and 3 respectively, and f_0 & f_1 are the focal distances of R and S from O respectively. Using the cosine formula for triangles

(with the external angle convention, section 1.5):

$$\begin{aligned} f_{0m} &= \left[f_0^2 + r_m^2 - 2f_0 r_m \cos\left(\frac{\pi}{2} + \theta_0\right)\right]^{\frac{1}{2}} \text{ ; and} \\ f_{1m} &= \left[f_1^2 + r_m^2 - 2f_1 r_m \cos\left(\frac{\pi}{2} - \theta_1\right)\right]^{\frac{1}{2}} , \end{aligned} \tag{4.26}$$

where r_m is the distance of point P_m from the origin and θ_0 and θ_1 are the angles of inclination of focal distances f_0 and f_1 from the horizontal respectively. Applying a binomial expansion of the form $(a+b)^n \approx a^n(1+\frac{nb}{a})$ to equations 4.26 above gives

$$\begin{aligned} f_{0m} &= f_0\left(1 + \frac{r_m^2 + 2f_0 r_m \sin\theta_0}{2f_0^2}\right) \text{ ; and} \\ f_{1m} &= f_1\left(1 + \frac{r_m^2 - 2f_1 r_m \sin\theta_1}{2f_1^2}\right) \end{aligned} \tag{4.27}$$

This approximation is akin to third order theory in geometric optics (refer to [HEC87, pages 133 and 221]), and so we should expect the same five primary Seidel aberrations to be present.

Substituting these expressions for f_0 and f_1 into equation 4.25 gives

$$\frac{n_1}{f_0} + \frac{n_3}{f_1} = \frac{m\lambda}{r_m^2} + \frac{2}{r_m}(n_3 \sin\theta_1 - n_1 \sin\theta_0) \tag{4.28}$$

This is an important result because between recording and replay the elliptical fringes (or zones) are fixed, so that for a given value of m, (a particular fringe), only λ, n_1, n_3, θ_0, f_0, θ_1 and f_1 are free variables. Thus changing λ between recording and replay will change the other free variables according to the relationship 4.28 above. There is a parallel expression for reflection gratings with focusing power, which can be derived in a similar way as

$$\frac{1}{f_0} + \frac{1}{f_1} = \frac{m\lambda}{n_1 r_m^2} - \frac{2}{r_m}(\sin\theta_1 + \sin\theta_0) \tag{4.29}$$

Both expressions 4.28 and 4.29 bear more than a passing resemblance to the thin lens equation and the Fresnel zone plate equation. They illustrate the inverse relationship between the focusing power of an HOE and the wavelength at which it is used. The

sign conventions for distances in front of and behind the HOE are the same as for the thin lens equation.

4.3.1 Calculating recording points

Using the notation of section 4.1.1 for wavelength-shifted planar gratings, equation 4.28 (and the corresponding reflection expression) can be used to calculate the recording points for an HOE recorded at λ_0 which is to be replayed at λ_1. In figure 4.7 the focal points R and S are at angles θ_0 and θ_1 to the horizontal; taking these angles as the average Bragg angles for the focusing HOE, the values of θ_0 and θ_1 can be calculated for the wavelength-shift. These angles, at λ_0, are θ_{00} and θ_{10} and can be calculated from equations 4.13 and 4.14 and applying Snells law. The zone plate style equation 4.28 can now be expressed as $n_1/f_0 + n_3/f_1 = k'(\lambda)$, where $k'(\lambda)$ is a dispersive constant given by equation 4.28. The values of m and r_m are constant between the two wavelengths for a wavelength-shift and can each be arbitrarily set to unity, so

$$\frac{n_1}{f_0} + \frac{n_3}{f_1} = \frac{1}{k(\lambda)} \,, \tag{4.30}$$

where $1/k(\lambda) = \lambda + 2(n_3 \sin\theta_1 - n_1 \sin\theta_0)$. Thus for a wavelength-shift between replay focal lengths f_{01}, f_{11} and recording focal lengths f_{00}, f_{10}:

$$k(\lambda_0)\left[\frac{n_{10}}{f_{00}} + \frac{n_{30}}{f_{10}}\right] = k(\lambda_1)\left[\frac{n_{11}}{f_{01}} + \frac{n_{31}}{f_{11}}\right] \,, \tag{4.31}$$

where

$$\begin{aligned} \frac{1}{k(\lambda_0)} &= \lambda_0 + 2(n_{30}\sin\theta_{10} - n_{10}\sin\theta_{00}) \text{ ; and} \\ \frac{1}{k(\lambda_1)} &= \lambda_1 + 2(n_{31}\sin\theta_{11} - n_{11}\sin\theta_{01}) \end{aligned} \tag{4.32}$$

A similar equation to 4.31 applies for reflection gratings, except the expressions of equation 4.32 are governed by equation 4.29. It can be seen from equation 4.31 that there are an infinite set of solutions for f_{00} and f_{10}. The three obvious sets to choose

are either:

- $f_{00} = \pm\infty$: implying a plane wave and point source recording geometry (circular fringes),
- $f_{00} = f_{10}$: a two equidistant point source recording geometry (elliptical fringes); and
- $f_{10} = \frac{\lambda_1}{\lambda_0} f_{11}$ and f_{00} from equation 4.31 (elliptical fringes): a point source recording geometry derived from the dispersive equation 4.22.

Note that the recording point sources need not be on separate sides of the HOE and it is advisable to use a ray tracing program, that can handle HOEs, to check the geometric spot produced by the geometry. In the eventuality that the solution given by equations 4.13, 4.14 and 4.31 do not provide an adequate geometric spot size (since the binomial approximation used in equation 4.27 can be inappropriate for low f/number lenses), then it gives a starting point for a ray tracing/spatial search approach such as that of Redmond [RED89a].

The model presented above seeks to optimise both the Bragg condition and replay spot size. If spot size is of paramount importance and efficiency can be sacrificed, then this method is still applicable, however the Bragg matching of fringe planes inside of the holographic medium can be ignored. In this case the recording angles are simply derived from the geometric grating equation 4.4, (relaxing the constraint 4.8), then inserted into equations 4.31 and 4.32.

Substrates are not included in the model and consequently were not included in the ray tracing used to check the model; the presence of a substrate simply increases the designed replay focal length by $\Delta f = (1 - n_{31}/n_s)d$, (or $(1 - n_{11}/n_s)d$, depending upon which side of the HOE the replay focal spot lies), where n_s is the substrate refractive index and d is the substrate thickness. It is therefore simple to change the design focal length to account for the substrate. If the replay focal length lies within the substrate then that can be included in the model by setting $n_{31} = n_s$. The inclusion in the model of a refractive index change in the holographic medium (between n_{20} and n_{21}) allows for thickness and index variations introduced by the processing procedures used.

4.4 Analysis of the method

As outlined in section 4.2, the quality of the focal spot is dependent upon factors other than the recording points. This section attempts to quantify the variation in spot size and Bragg mismatch as each of the recording variables change; being interlinked, it is impossible to isolate the effects of one variable from another. General trends can, however, be found.

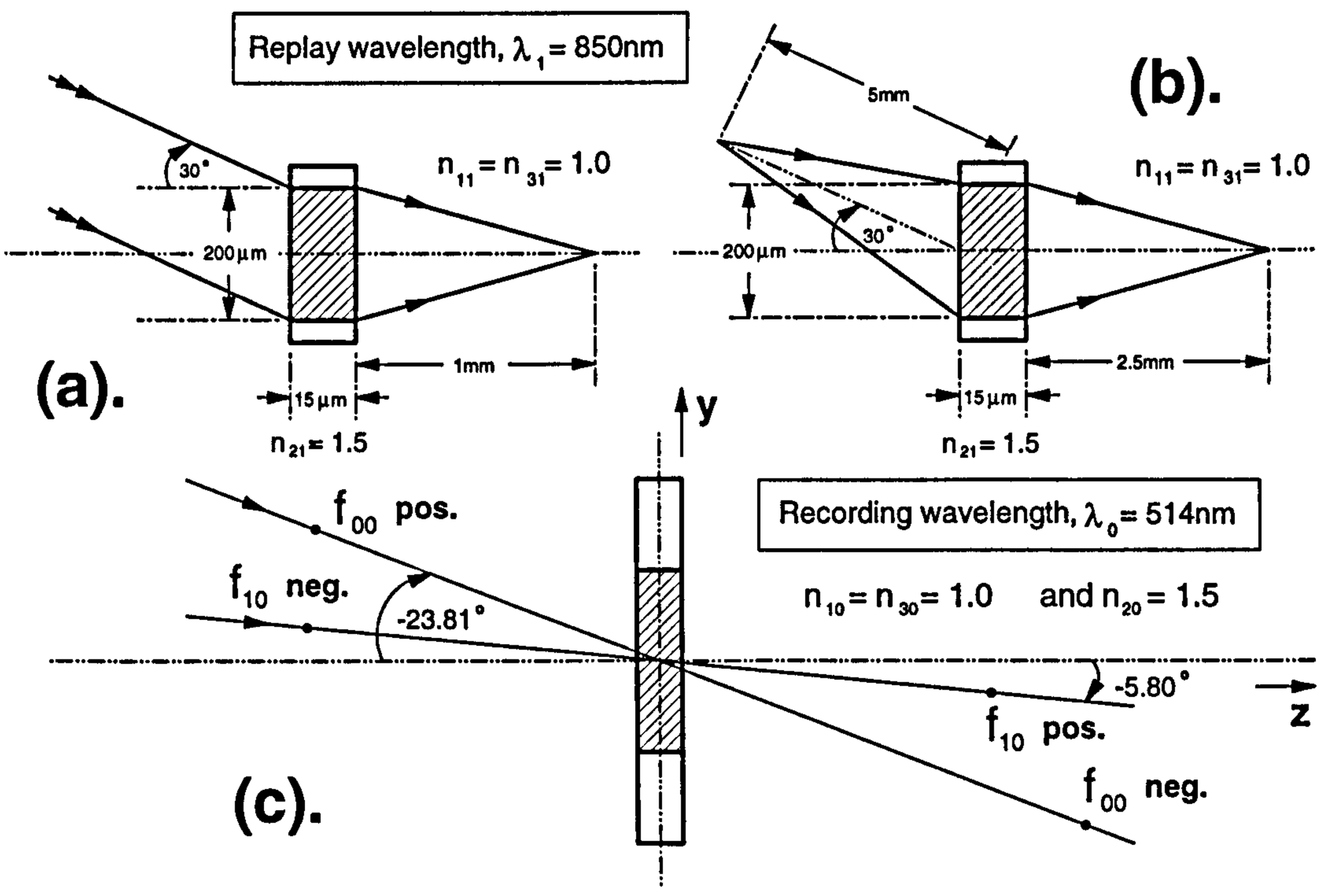


Figure 4.8: The replay and recording geometries for the two HOEs considered in the method analysis: (a). Replay for plane wave to focal point, (b). Replay for point to point; and (c). Recording geometry for both replay geometries (same recording angles for both). Recording focal lengths f_{00} and f_{10} can be either positive or negative - either side of the HOE, they have no fixed basic values since the method is designed to determine these, (see section 4.4.1).

The two replay geometries shown in figure 4.8 were chosen to investigate the effects of these variables. They represent the two most useful lenslet configurations used throughout the interconnect work in this thesis:

- A plane wave incident off-axis being brought to a focal point on-axis (figure 4.8a); and
- A focal point off-axis being brought to a focal point on-axis (figure 4.8b).

From appendix B it is clear that these represent circular and elliptical surface fringe profiles respectively. The values for the variable parameters shown in figure 4.8 are assumed as base levels, (recording points do not have base levels - optimum values are determined from the analysis in section 4.4.1). In each analysis, (apart from that of recording points) only one of these values is allowed to vary from the base level. Thus, the wavelength shift is always from 514.5nm to 850nm unless the parameter being analysed is the magnitude of the wavelength shift. This wavelength shift applies to both geometries and so the angles of the recording beams with respect to the hologram will be the same in both instances, (the values shown in figure 4.8c are derived from the average Bragg condition at the centre of the HOE as specified in section 4.1.1).

Analysis criteria

The principle tool used to analyse this method of design is a ray tracing program which is capable of dealing with volume HOEs. This was developed by Redmond, (the exact details can be found in [RED89a, pages 184-206], but a brief description is included in section 4.2.1) as an extension of the ray tracing techniques of Welford [WEL74] and Latta [LAT71b]. Optical Research Associates (ORA) CodeV ray tracing package [ORA94] was also used to verify the results. For each variable change 120 rays were traced to find the spot size and the Bragg mismatch, (more than this does not significantly improve the accuracy of the analysis).

The ray tracing program takes worst case rays for both the spot size and the Bragg mismatch calculations. Spot sizes are defined by the location of the farthest rays each side of the spot centre. The Bragg mismatch is calculated as the maximum (worst case) angular difference between the incident ray and the Bragg angle (calculated from the recording points) for each sector of the hologram being traced. The worst case spot sizes are not always as bad as they seem, for most cases where the traced spot size is significantly larger than the diffraction limit it is found that 90% of the rays usually lie in a spot area 55% to 70% smaller. The worst case Bragg error is, however, absolute because it represents an efficiency variation across the HOE.

The spot sizes produced by the ray tracing are $x \times y$ microns, to graph these sizes it is more convenient to have a single number. There are many ways to represent the spot size as a single number, the one adopted is a root mean squared average of the form $\sqrt{\frac{x^2+y^2}{2}}$. This produces a mean value when $x \simeq y$ and takes adequate account of the difference when $x < y$ or $x > y$. Since the power density in a focal spot will ultimately be limited by diffraction, the effect on the spot size of a geometric aberration will be negligible for aberrations less than the diffraction limit. The final focal spot (point spread function) produced by a lens results from the convolution of the aperture function of the lens (diffraction limit) with the geometrical image defects (aberrations) [HEC87, pages 485-486]. However, an estimate of the real spot diameter, d_s, to be expected in terms of the geometric (d_g) and diffraction-limited ($d_d = {}^{2f\lambda}/_{a}$, where a is the aperture side of the square HOEs in the analysis) spot diameters is:

$$d_s = \sqrt{d_g^2 + d_d^2} \tag{4.33}$$

The spot size is of course not the only condition to be satisfied in order for a lens to fulfill the diffraction-limited criteria. The other is that the measured irradiance profile of the focal spot should closely fit the $sinc^2$ function for the lens [BOR91]. This criterion is not easily calculated and as such is not represented in this analysis.

For each spot size produced in the following analysis there are two different values quoted: the best focal spot and the designed focal spot. The designed spot is that traced to produce a focal spot at the exact focus of the replay geometry. The best focal spot is (usually) smaller than the designed spot and is the result of a focal plane scan by the ray tracing program; it is the minimum calculated blur circle. Where the best spot appears larger than the design spot this is due to the "aspect ratio" of the spot; the focal plane scan usually picks a blur circle in which $x \simeq y$, (for an $x \times y$ spot). The design focal length is usually close to the best focal length, the mismatch varies most significantly with the f/number of the lens. The worst is about a 15% variation but more usually it is around 1%. The mismatch is usually well within the Rayleigh range of the beams manipulated by a specific interconnect (for a Gaussian this is given by, $z_R = \pi\omega_o^2/\lambda$, where ω_o is the beam waist radius).

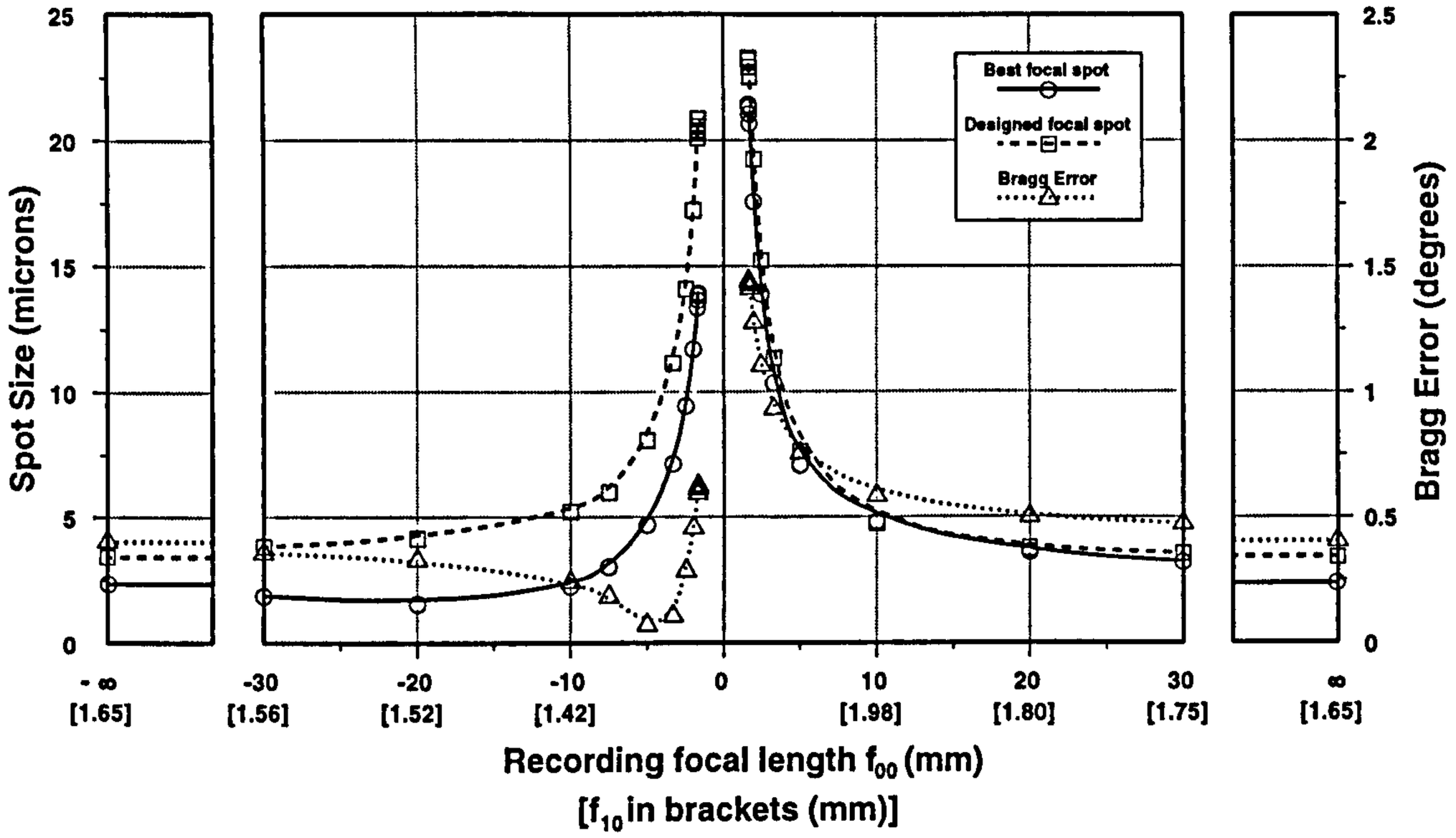


Figure 4.9: Plane wave to focal point replay geometry spot size and Bragg error variation with recording points fulfilling the average Bragg condition and obeying equation 4.31.

Where this is not the case a small optimisation of the recording points can be carried out (using Redmonds method) with this ray tracing program to produce a minimum blur circle at the design focus.

The first optimisation is that of the best recording points for the two geometries. The points are then fixed in the analysis of other parameter variations.

4.4.1 Recording points

The basic parameters for replay and recording of the HOEs are shown in figure 4.8. When the base level parameters are applied to equations 4.31 and 4.32 a value for $k(\lambda)$ can be found for both cases, and used to derive values for the recording points, f_{00} and f_{10}:

$$\begin{aligned} \frac{1}{k(\lambda)} &= \frac{k(\lambda_1)}{k(\lambda_0)}\left(\frac{1}{f_{01}}+\frac{1}{f_{11}}\right) \text{ ; and} \\ \frac{1}{f_{00}}+\frac{1}{f_{10}} &= \frac{1}{k(\lambda)} \end{aligned} \tag{4.34}$$

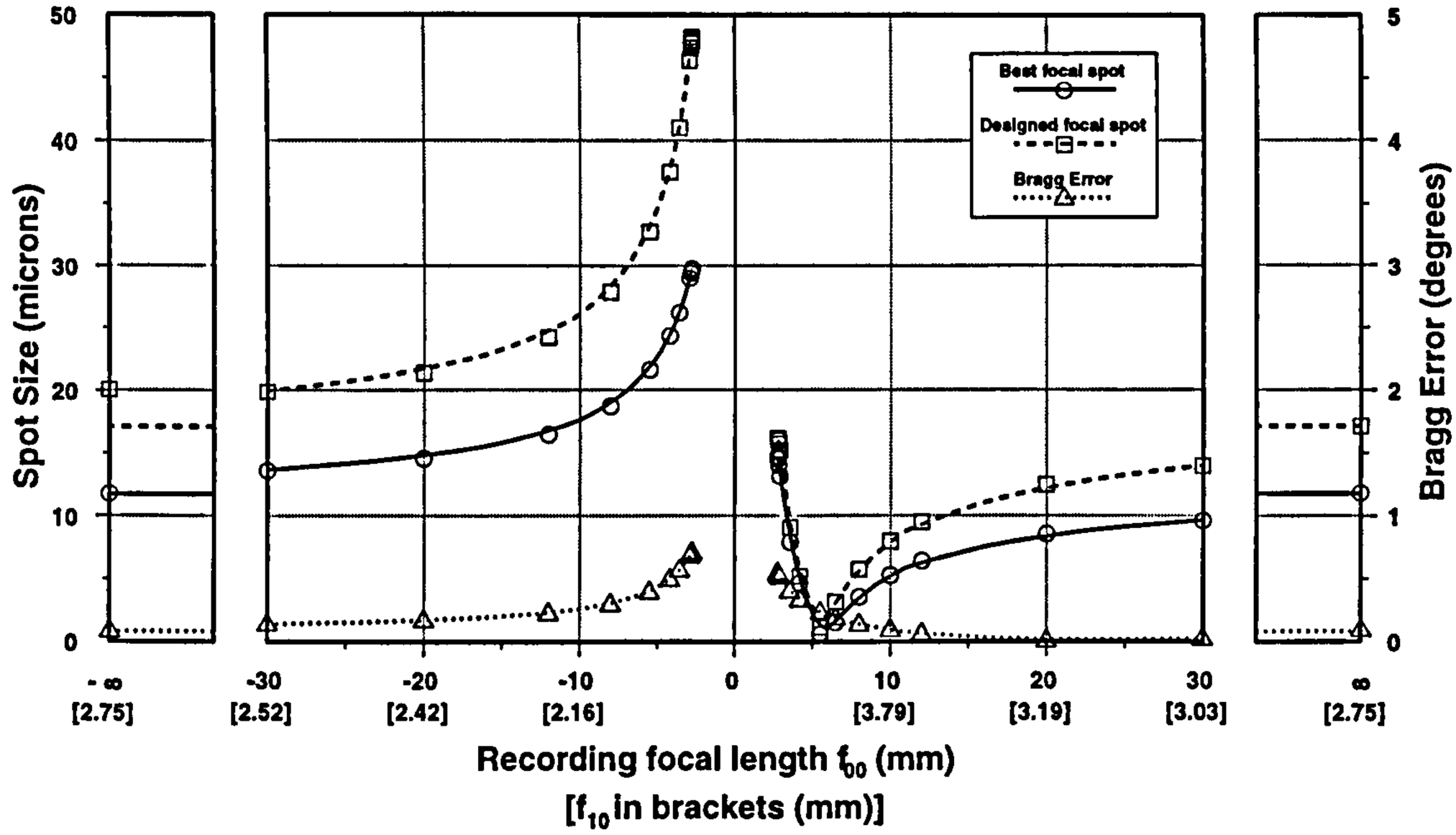


Figure 4.10: Point to point replay geometry spot size and Bragg error variation with recording points fulfilling the average Bragg condition and obeying equation 4.31.

The values of $k(\lambda)$ for the replay geometries in figures 4.8a & b are 1.65 and 2.75 respectively. These values are used to calculate the dependence of spot-size on recording geometry, shown in figures 4.9 and 4.10.

- **Plane wave to focal point replay.** It is clear from figure 4.9 that a recording geometry with a plane wave, ($f_{00} = \pm\infty$) and a focal point, ($f_{10} = k(\lambda) = 1.65$mm) results in a geometric spot size of 2.5μm and a Bragg error of 0.41°. This is well below the diffraction limit of 8.5μm and the Bragg error is so small that it will have no noticeable effect on the HOEs efficiency. Although the empirical ray trace shows that recording points with $f_{00} = -20$ and $f_{10} = 1.52$mm, produce the smallest spot, the spot size differs only marginally from the infinity limit (these were used as base values in the analysis of other variables). It is generally true from other replay designs conforming to this same basic geometry that the smallest spot size is achieved when:

$$\begin{aligned} f_{00} &= \pm\infty \text{ ; and} \\ f_{10} &= \frac{k(\lambda_1)}{k(\lambda_0)} f_{11} \end{aligned} \tag{4.35}$$

- **Point to point replay.** It is not exactly clear from figure 4.10 what are the optimum recording points; obviously a plane wave does not correspond to either recording beam in this case. The other two conditions which are noticeable are, from equation 4.34, $f_{00} = f_{10} = 2k(\lambda)$ and, from equation 4.22, $f_{10} = \frac{\lambda_1}{\lambda_0} f_{11}$. Figure 4.10 conforms almost exactly to the first condition. However, from several other similar geometries tried with slightly different designs the second condition is held to be more generally true. In either case the spot size is very similar and well below the diffraction limit of 21.3μm. The best recording points of $f_{00} = f_{10}$, producing a spot size of 0.7μm (and a Bragg error of 0.23°), are used as base values in the further analyses. However, the conditions which are more generally applicable in defining the best recording points for this type of replay geometry are:

$$\begin{aligned} f_{00} &= \frac{k(\lambda) f_{10}}{f_{10} - k(\lambda)} \text{ ; and} \\ f_{10} &= \frac{\lambda_1}{\lambda_0} f_{11} \end{aligned} \tag{4.36}$$

In the case of both replay geometries the general conditions for minimum geometric spot size do not correspond to the conditions for minimum Bragg error. This is unsurprising from the discussion in previous sections; this method satisfies the average Bragg angle at the centre of the hologram and so the difference between the minimum Bragg error and the Bragg error at the optimum recording points is negligible.

4.4.2 Replay wavelength

The base level values are used for both replay geometries, the recording points are taken from section 4.4.1, and the replay wavelength is now the variable parameter. The results of this analysis are shown in figures 4.11 and 4.12 together with the diffraction limit in both cases. It is clear that as long as the wavelength shift is to a longer wavelength there is no dramatic effect on the resultant spot size. If the shift is to a shorter wavelength the spot size very quickly rises above the diffraction limit,

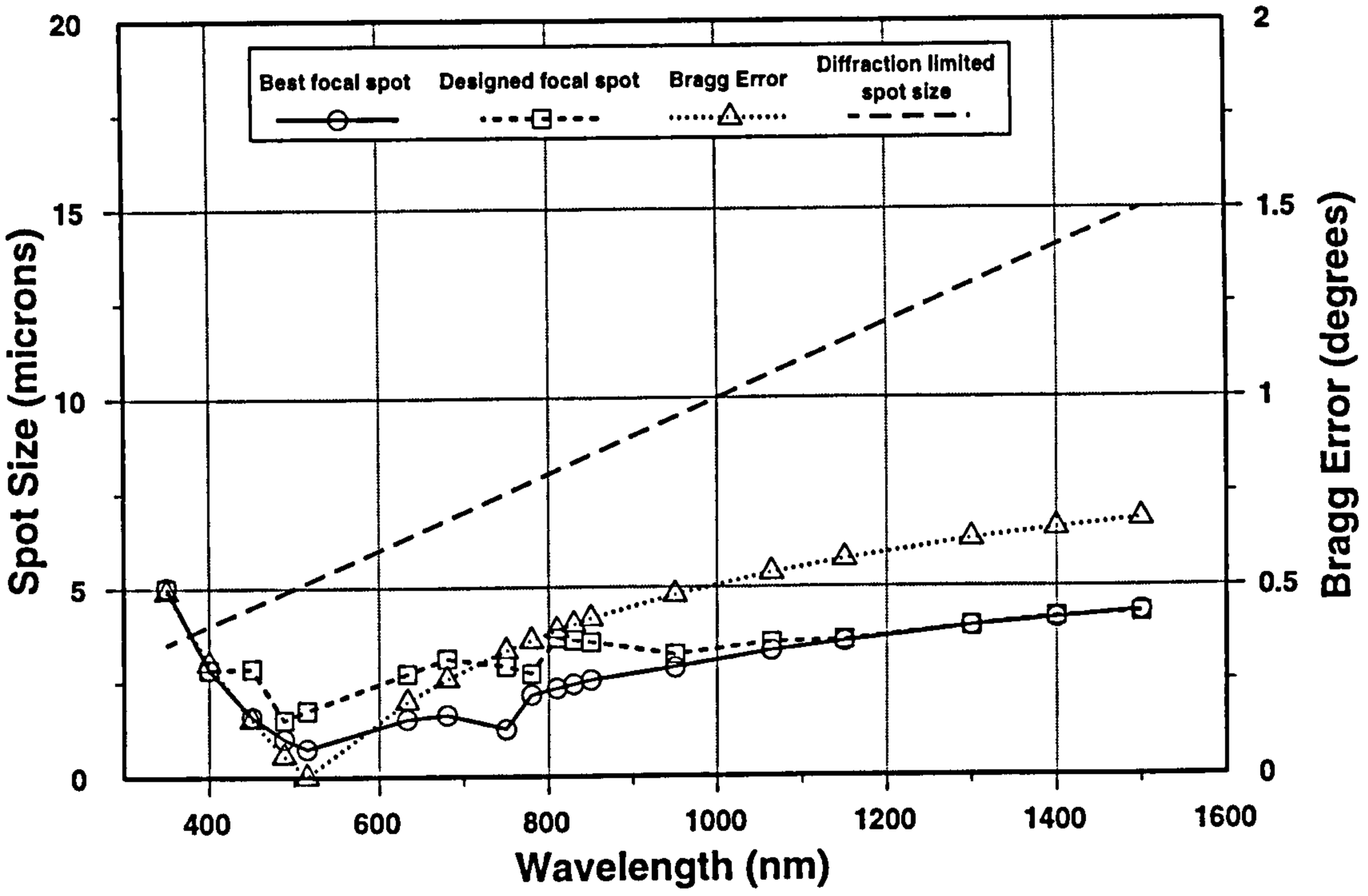


Figure 4.11: Plane wave to focal point replay geometry spot size and Bragg error variation with replay wavelength. The recording wavelength is 514nm.

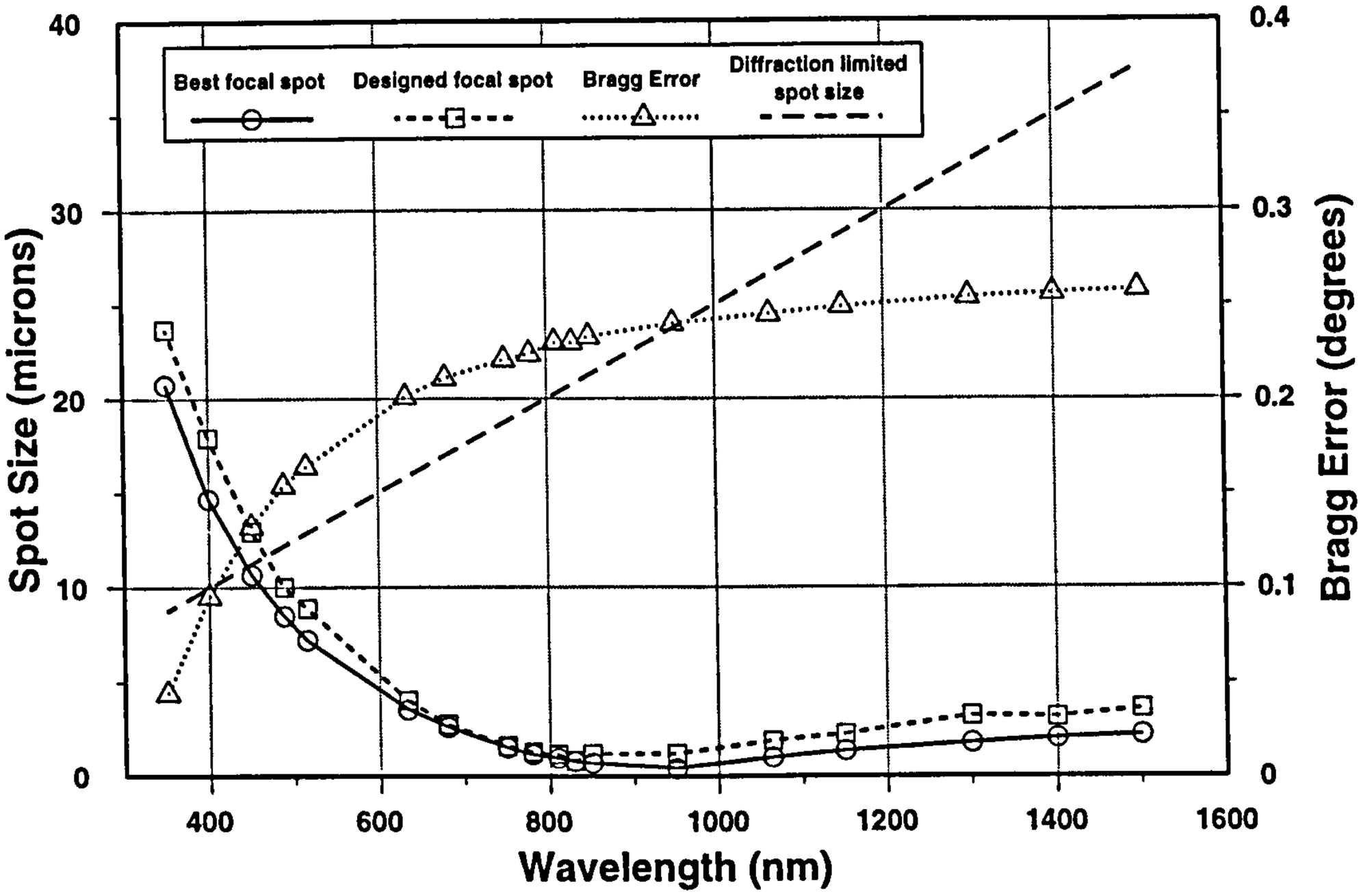


Figure 4.12: Point to point replay geometry spot size and Bragg error variation with replay wavelength. The recording wavelength is 514nm.

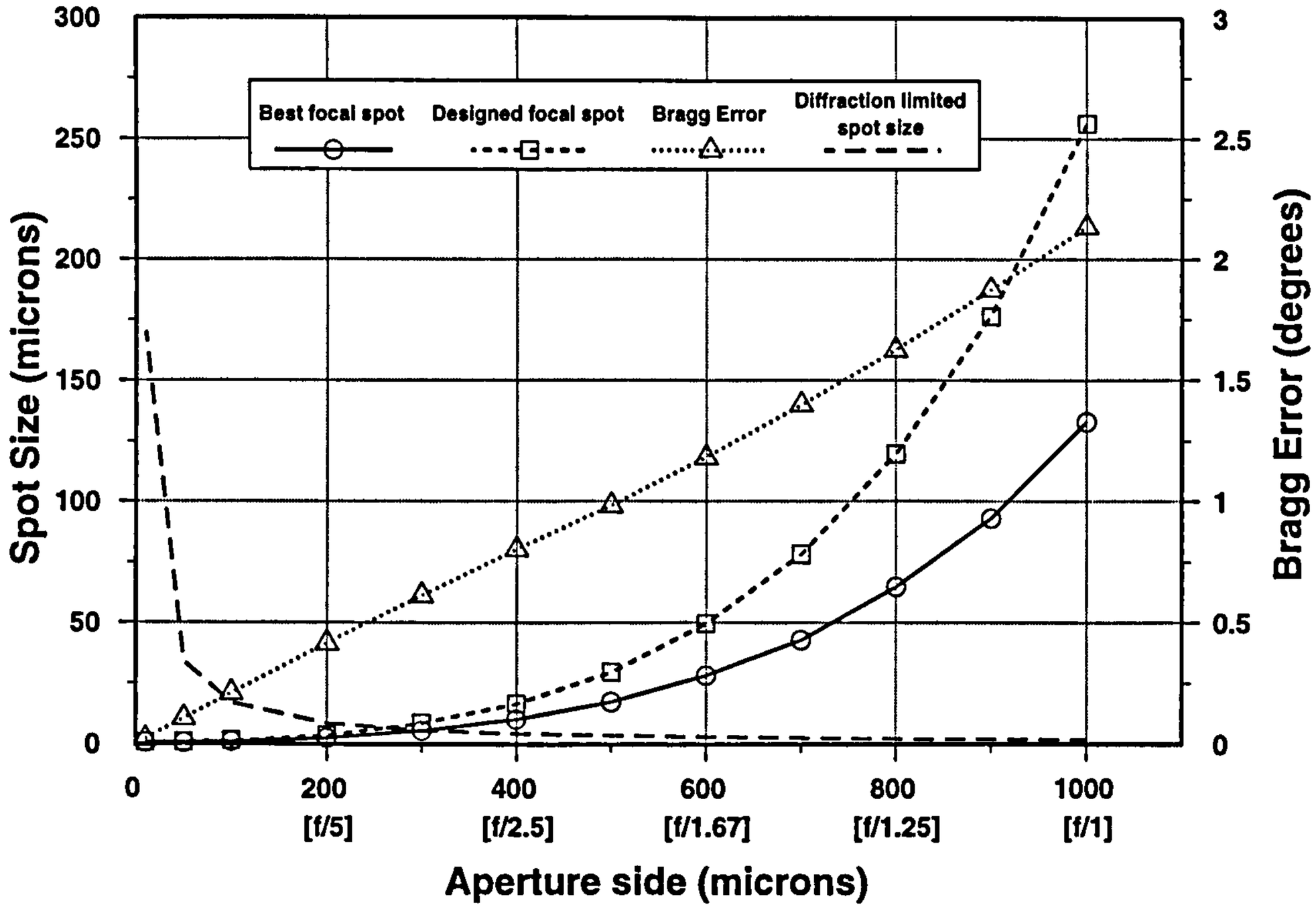


Figure 4.13: Plane wave to focal point replay geometry spot size and Bragg error variation with hologram aperture size; at a fixed replay focal length. The f/number of the HOE on replay is shown in brackets.

this is born out by analysing the dispersive aberration equation 4.24. Since DCG is sensitive in the blue and green this does not present a problem, because a shorter wavelength shift is not required.

As the wavelength shift increases, the Bragg error in both geometries increases up to a limit that is always under 1° and does not represent a significant efficiency variation.

4.4.3 Aperture size at fixed focal length

The aperture size is the variable parameter here, and all other variables assume the base level. From figures 4.13 and 4.14 it is clear that as the aperture side increases, (reducing the effective f/number of the lenslet), the spot size becomes dramatically worse. The break even point, at which the aberration is the same size as the diffraction limit, is at about $f/3$ for both geometries. At $f/1$, the spot sizes are comparable to the aperture and are therefore likely to be of little use. These

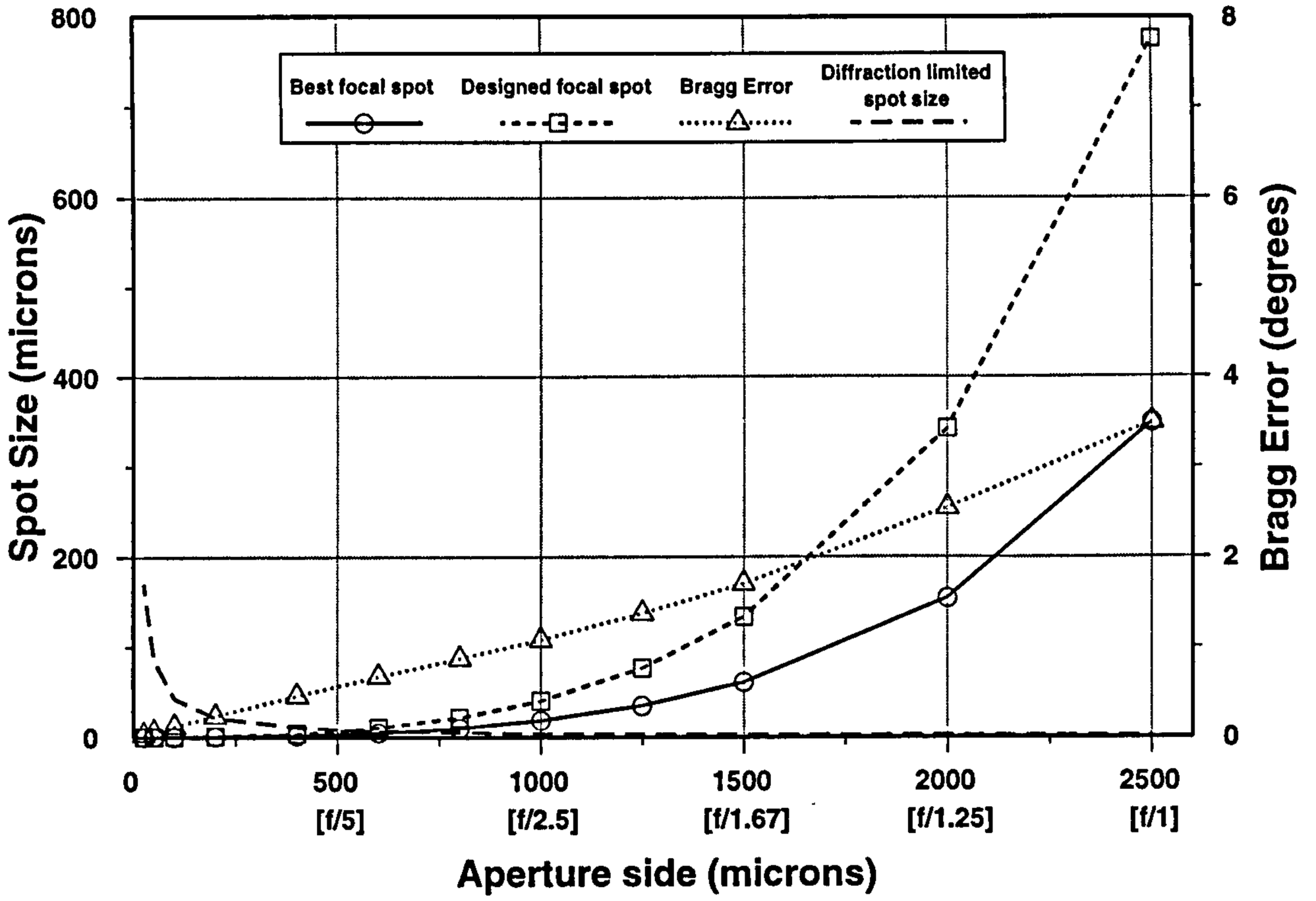


Figure 4.14: Point to point replay geometry spot size and Bragg error variation with hologram aperture size; at a fixed replay focal length. The f/number of the HOE on replay is shown in brackets.

are, however, worst case rays and in fact 90% of the spots are actually contained in an area 60% to 70% smaller than the worst case. The average Bragg angle does get worse, at lower f/numbers, as the marginal rays depart further from the average Bragg condition at the centre of the lens.

4.4.4 Focal length at fixed aperture

Not surprisingly, the result of a focal length variation is the inverse of an aperture size variation, (at a given f/number). This can be seen quite clearly from figures 4.15 and 4.16, again the break even point is around $f/3$ in both cases. The spot sizes at smaller f/numbers are better than for the corresponding values with aperture variation, but this is due to the HOE being smaller, (since aberrations scale linearly with physical size). The worst case rays again belie the 90% power levels which are routinely 30% to 45% of the sizes graphed. Again, the average Bragg error is considerably worse at low f/numbers.

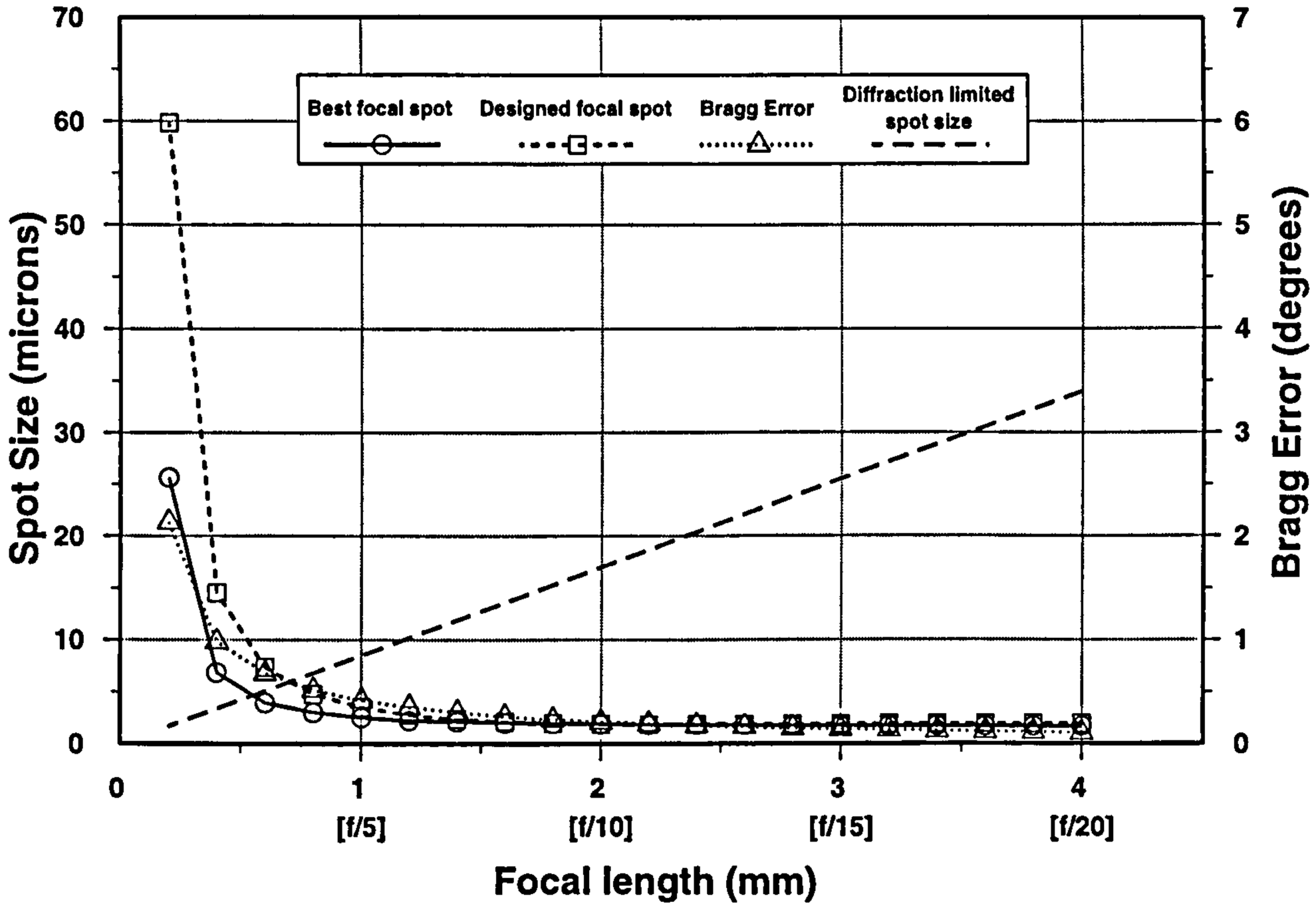


Figure 4.15: Plane wave to focal point replay geometry spot size and Bragg error variation with replay focal length; at a fixed aperture size. The f/number of the HOE on replay is shown in brackets.

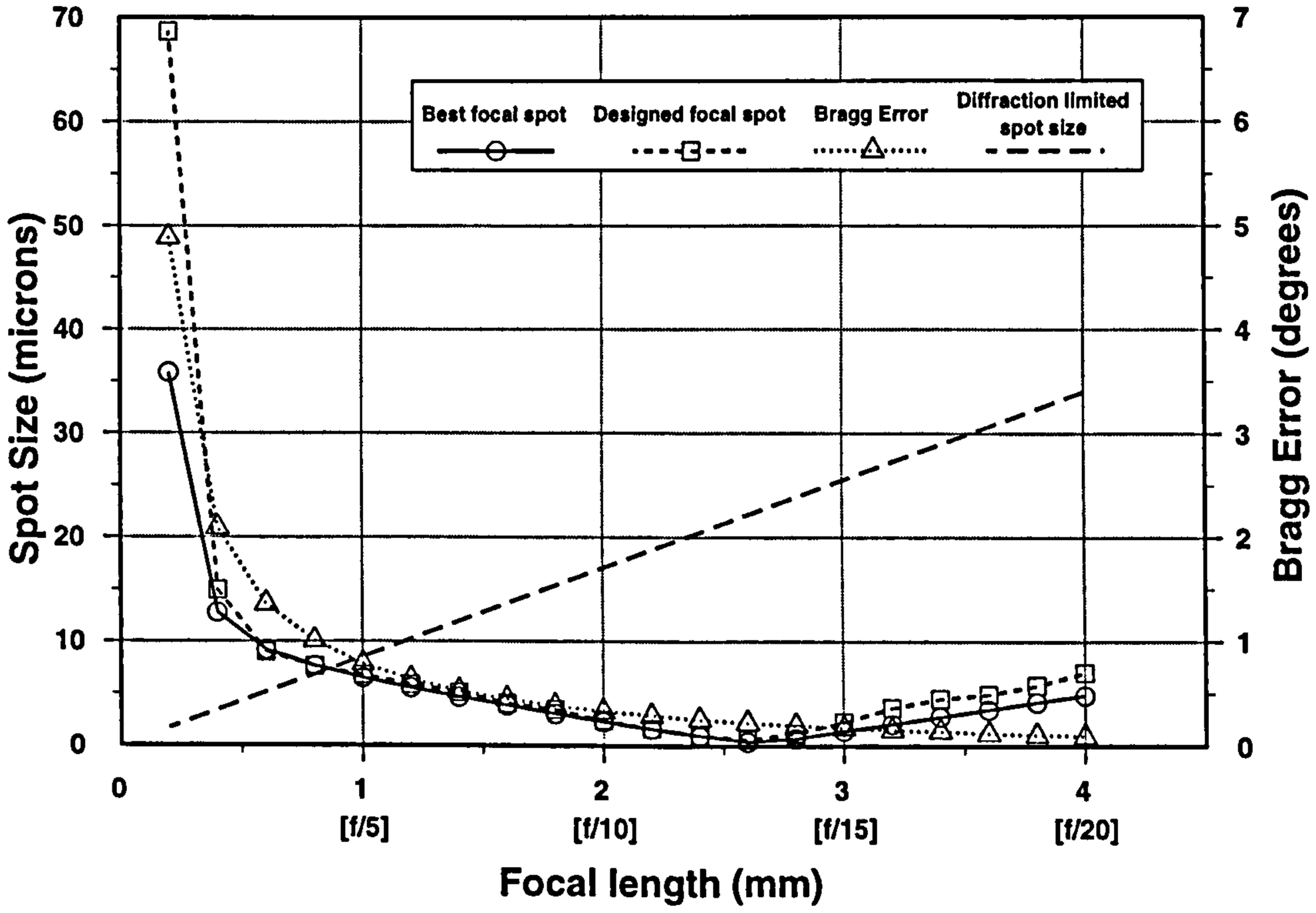


Figure 4.16: Point to point replay geometry spot size and Bragg error variation with replay focal length; at a fixed aperture size. The f/number of the HOE on replay is shown in brackets.

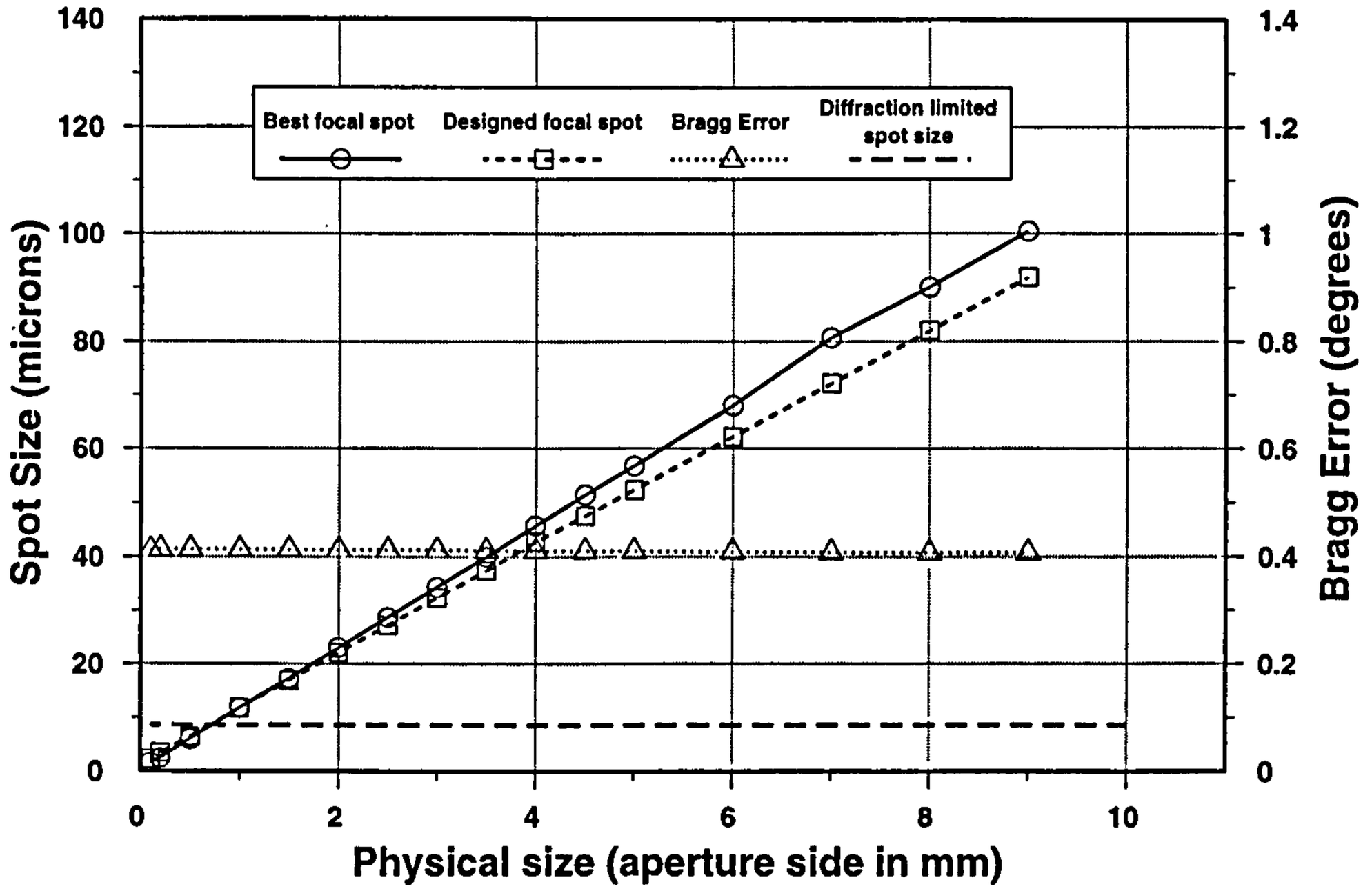


Figure 4.17: Plane wave to focal point replay geometry spot size and Bragg error variation with physical HOE size; at a constant replay f/number.

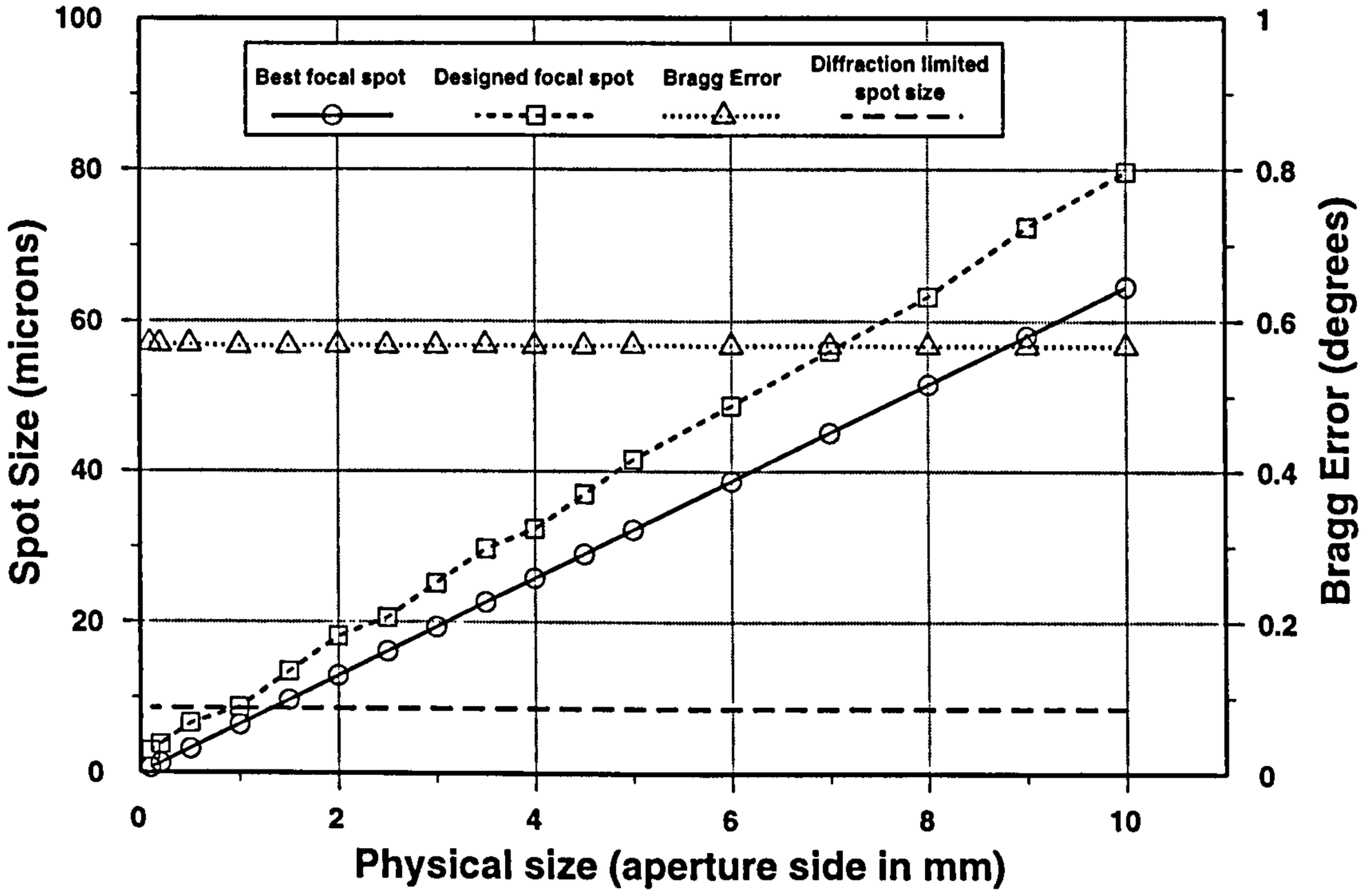


Figure 4.18: Point to point replay geometry spot size and Bragg error variation with physical HOE size; at a constant replay f/number.

4.4.5 Physical size

The replay geometries are changed slightly here, the base level f/numbers of $f/5$ and $f/12.5$ being both set to $f/5$ for this analysis, (this is done to monitor the effects of a physical size change at the same f/number for the two replay geometries). The focal lengths and aperture sides are varied to maintain $f/5$ as the physical size of the holograms increases. Figures 4.17 and 4.18 show that the aberrations increase proportionally with the physical size in both instances, the diffraction limit crossover point occurring at about 750μm. A larger f/number will increase this physical size limit. The spots simply magnify without changing shape, the spot size at 1cm being twice that at 0.5cm. Unlike an aperture variation, spot sizes cannot be effectively reduced by worst case rays. The Bragg error is constant in both cases for fixed f/number geometries.

4.4.6 Replay angle & emulsion thickness

The effects, on spot size and Bragg error, of varying the replay angle (whilst maintaining the average Bragg condition) and changing the hologram thickness are not significant. The contribution to aberrations of replay angles in the range 10° to 60° and emulsions in the range 3μm to 60μm are all well below the diffraction limit for both geometries, (consequently graphs are not presented).

4.4.7 Conclusions from the analysis

This analysis is carried out on transmission gratings, because they are the main HOEs of interest for interconnects with focusing power in this thesis. The method has not been analysed for reflection gratings with focusing power. However, since reflection holograms can be fabricated to have a large angular bandwidth (see section 6.2), they could prove to be less demanding on the Bragg matching condition. This could consequently be relaxed in favour of geometric aberration minimisation.

The conclusions on the usefulness of the design method drawn from this analysis may be presented as follows:

- **Spot size.**

  1. The best recording conditions are summarised in section 4.4.1.
  2. With the exception of recording points, point to point and plane wave replay geometries behave similarly. Further testing of the two geometries (not presented in the preceding analysis) where the replay focal spot was off-axis produced markedly poorer spot sizes (and larger Bragg errors).
  3. Spot sizes quoted are for *worst case* rays. Consequently, where focal spots appear to be significantly larger than the diffraction limit, 90% of rays traced can fall in a spot area upto 70% smaller than the worst case figure.
  4. The parameters that have the most marked effect on the spot size are:
     - f/number. With decreasing f/number, the point at which the geometric spot size exceeds the diffraction limit is about $f/3$.
     - Physical size. The physical size limit (where geometric aberrations equal the diffraction-limited spot size) for $f/5$ lenses at 850nm is 750μm, this increases as the f/number increases and vice versa; for the $f/3$ limit (above) the physical size limit is (from ray tracing) 300μm.

  Considered individually, the other replay conditions have a relatively small effects on the spot size, however, the cumulative effects of these aberrations can be significant. The aberrations introduced by the variation in these other parameters increases as f/number decreases, (or as physical size increases).

- **Bragg error.**

  1. The worst ray Bragg error is approximately inversely proportional to the f/number of the lens.
  2. It increases up to a limit with the magnitude of the wavelength shift. Although this limit is not large for the test geometries it scales approximately inversely with f/number.

3. The other parameters do not affect the size of the Bragg Error.

This model produces results which are generally in agreement with the method of Redmond [RED89a] in terms of recording points that are generated and replay spot sizes. In addition, the recording points are in agreement with those designed by ORAs CodeV (in all instances tested) and replay spot sizes are very similar to the results of Falkenstorfer et al. [FAL91], a generalised ray tracing method similar to Redmond. It would seem clear that this simple technique represents the best that can be achieved using spherical recording wavefronts. Improvements to the spot size aberrations can only be made by resorting to the more complex recording schemes involving CGH or ancestral volume HOEs to produce aspheric wavefronts. Although aspherics will improve the spot sizes at small f/numbers, the Bragg error will be largely unchanged because of the trade-off between Bragg mismatch and spot size. Bragg errors in excess of 1° are often large enough to cause an excessive efficiency limitation and this is not readily overcome without sacrificing spot size. This problem could in some cases be overcome by using the large angular bandwidth exhibited by reflection gratings.

4.5 Volume HOE efficiency considerations

A holographic lenslet, designed using the method of section 4.3, will diffract most efficiently at the average Bragg angle. Consequently the cone of light being focused by a lenslet will not all be uniformly redirected into a diffraction-limited spot; there will be a deviation from the ideal maximum at the extremities of the aperture of the HOE, represented by the Bragg error in section 4.4. The optimum exposure for the maximum efficiency of a holographic lens, being replayed at the wavelength at which it was recorded, can easily be ascertained by recording a test plate of HOEs with a range of exposures. The same procedure can be applied to wavelength-shifted gratings where a coherent light source is available at the replay wavelength. The method outlined here is used to gauge the optimum exposure for maximum diffraction efficiency of wavelength-shifted gratings where a replay source is not

readily available at the new wavelength.

4.5.1 Wavelength-shifted efficiency optimisation

The following optimisation procedure applies equally well to planar gratings as it does to holographic lenses:

- Record a test plate of gratings at λ_0 over a range of exposures, in the wavelength shifted geometry. Bear in mind that the index modulation required to achieve peak diffraction efficiency at the replay wavelength, λ_1, must be approximately λ_1/λ_0 times greater than that at λ_0.

- Measure and note the diffraction efficiency, of the resultant test plate HOEs, for each exposure level.

- Using a volume hologram analysis tool, such as the rigorous coupled wave theory of chapter 3, deduce the refractive index modulation for each of the diffraction efficiencies measured from the test plate; providing a graph of index modulation versus exposure energy.

- After using Fourier analysis on the expression for refractive index modulation $\Delta n(x)$, from section 3.8.1:

$$\Delta n(x) = \Delta n_{max}\left[1 - \exp\left(-\frac{\beta E}{\Delta n_{max}}(1+\cos Kx)\right)\right], \qquad (4.37)$$

  find values of the variables, β and Δn_{max} that fit the experimental data of index modulation and exposure energy.

- Using the volume hologram analysis tool find the index modulation that gives the maximum diffraction efficiency at λ_1 and the designed replay geometry .

- From the results of the Fourier analysis (a set of index modulation and exposure energies at λ_0), select the correct exposure at λ_0 to give the index modulation required for optimum diffraction at λ_1.

4.6 Experimental results

During the course of the research described in this thesis, many spherical and cylindrical wavelength-shifted lenses were recorded in DCG, using this method and that of Redmond. They have been used in a variety of communication and computational system experiments, working at a range of near IR wavelengths and having f/numbers from $f/4$ to $f/25$. The lens described here is typical of those recorded and embodies most of the difficulties encountered when making such holographic microlenses.

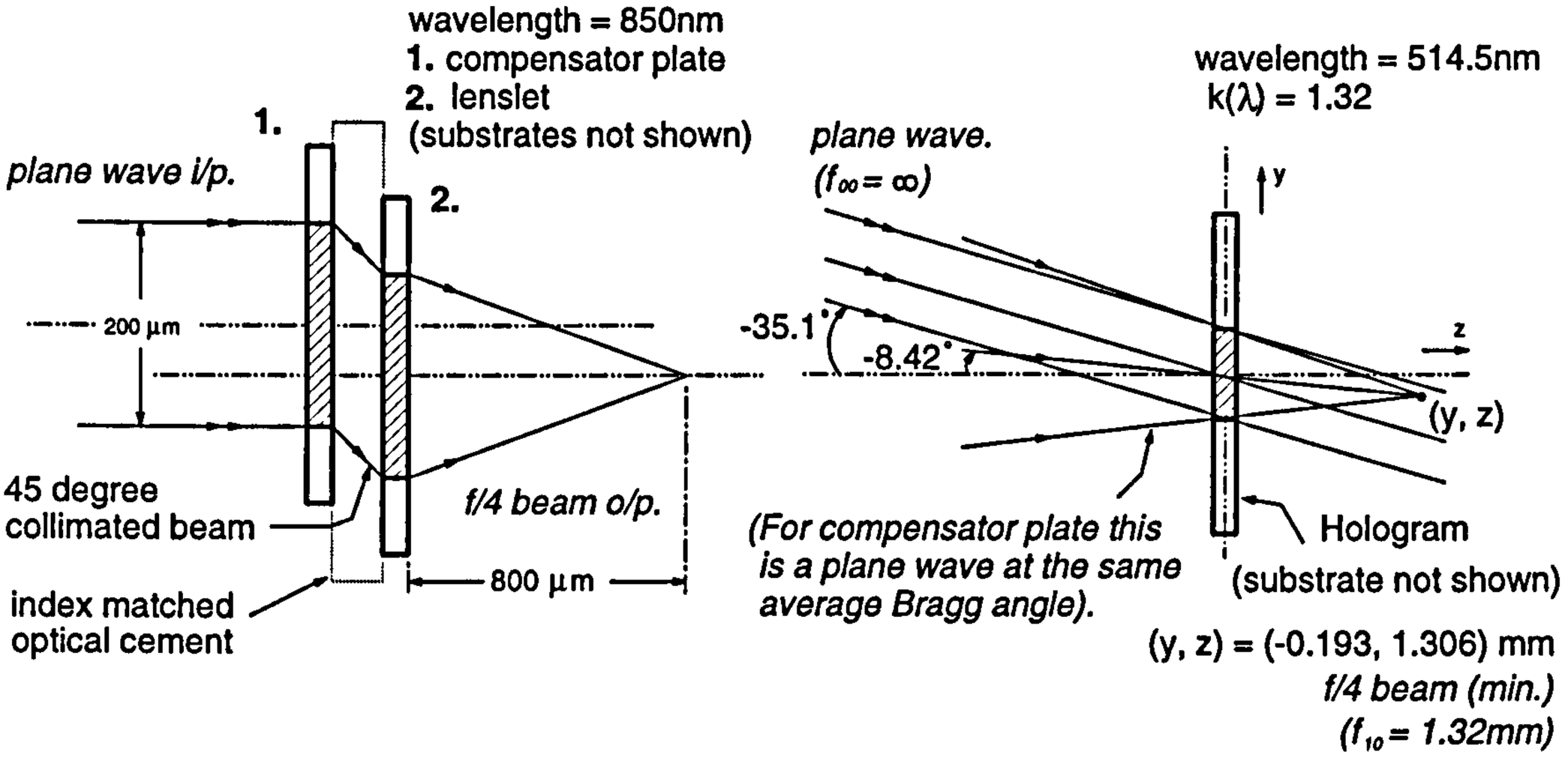


Figure 4.19: Cylindrical lenslet array: (a). On-axis replay geometry (with compensator plate); and (b). Recording geometry at 514.5nm (plane waves used at these angles to record compensator).

The lens replay geometry is shown in figure 4.19a and was designed for use in a demonstration optical crossbar working at 850nm utilising a 1×16 lens array described in section 6.3. The lens elements were rectangular of side $200\mu m \times 4mm$ and produced cylindrical wavefronts, resulting in a 4mm line focus. The lens had to work on-axis and in order to achieve this a 'compensator' plate was incorporated. This took a plane beam on-axis and simply redirected it to an angle which, in use, corresponded to the off-axis plane wave input to the lenslet. In this instance the angle was chosen to be 45°, this gives a longer optical path through the gelatin and hence a larger effective index modulation than a shallower angle of say 30°. The lens

array and compensator plate were cemented together, (gelatin-to-gelatin), to form a holographic doublet that produced a highly efficient on-axis diffractive lens. The optical system the lens was designed to be used in worked at $f/4$ and so the lens was designed to work in air with a focal length, f_{11}, of 800μm. These operational parameters led to the recording geometry shown in figure 4.19b, when the method described in section 4.3 was applied. This geometry was achieved on the optical bench using the recording set-up shown in figure 4.20. The recording arrangement produced both the wavefronts required for fabrication and clearly imaged apertures to achieve well defined microlenses that were closely packed, with minimum wastage and overlap.

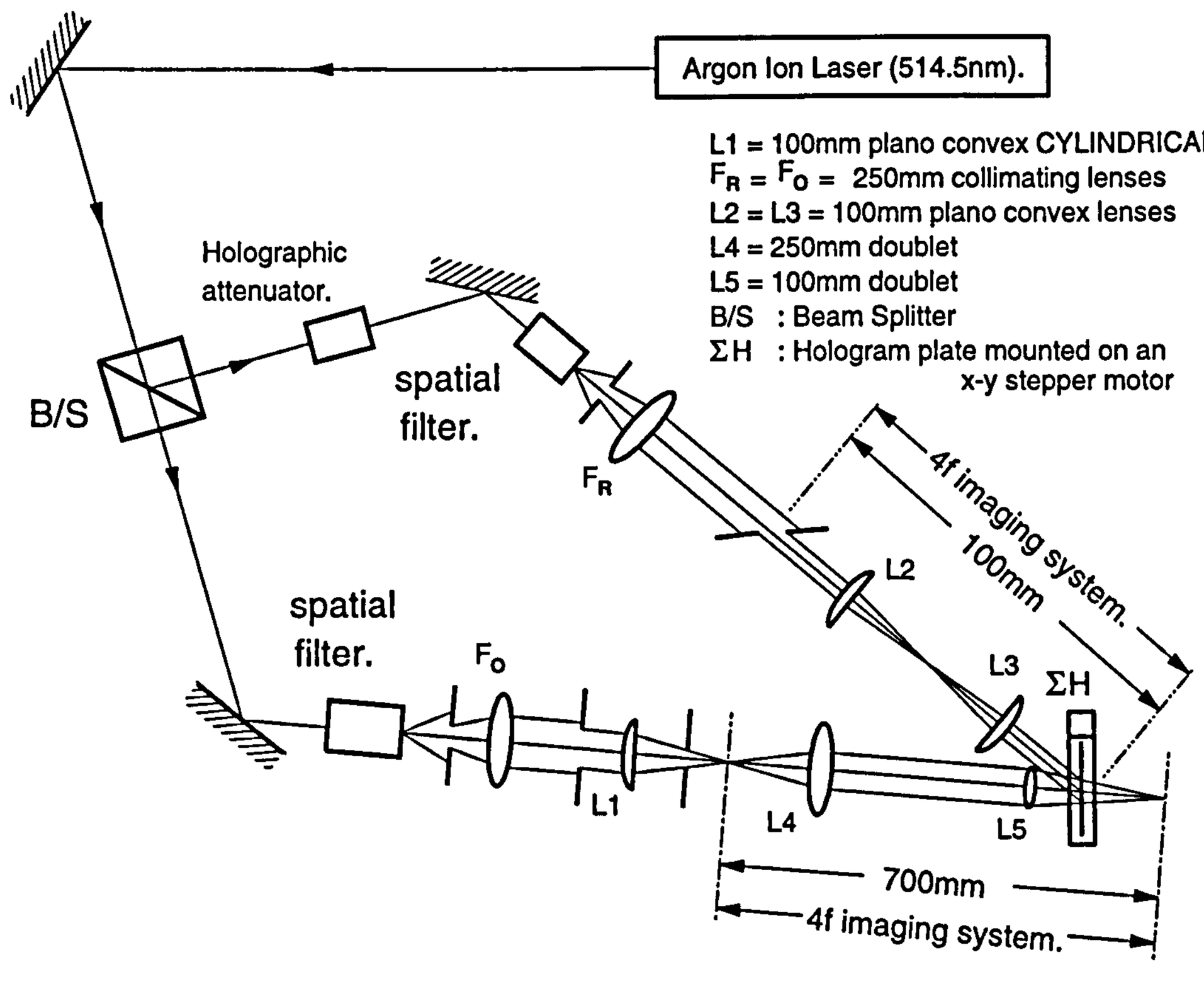


Figure 4.20: Recording arrangement at 514.5nm required to produce the recording geometry shown in figure 4.19b.

The diffraction-limited wavefront used to record the lens was equal in f/number to the desired replay wavefront; to this end, the cylindrical lens shown in figure 4.20 was used at $f/10$, (to minimise aberrations produced by this lens), and then demagnified through a $4f$ imaging system to achieve $f/4$ at 514.5nm.

The lens was designed to be cylindrical, but the ray traced spot sizes shown in figure 4.21 are for a lens which is spherical. This is because the ray tracing program is only designed to work with spherical lenses, although the design method is equally applicable to both, (Syms and Solymar discuss the replay of cylindrical holographic lenses [SYM82] at the recording wavelength). CodeV was used to check the design when applied to a cylindrical lens; the spot size predicted by CodeV is about 2.2μm, compared with the best spot size of 4.5μm and the design focus spot size of 5.4μm. As explained already the ray tracing program produces a conservative estimate of spot size based on worst case rays, whereas CodeV is an RMS spot diameter. The diffraction-limited spot size for this lens is 6.8μm, so from equation 4.33 we would expect the actual spot to be of the order of $\sqrt{6.8^2 + 2.2^2} = 7.2\mu$m. Figure 4.22b shows a CodeV point spread function (calculated using the recording points), being a convolved function it represents a more accurate spot size (7.98μm) for diffraction by the aperture and aberrations of the lens.

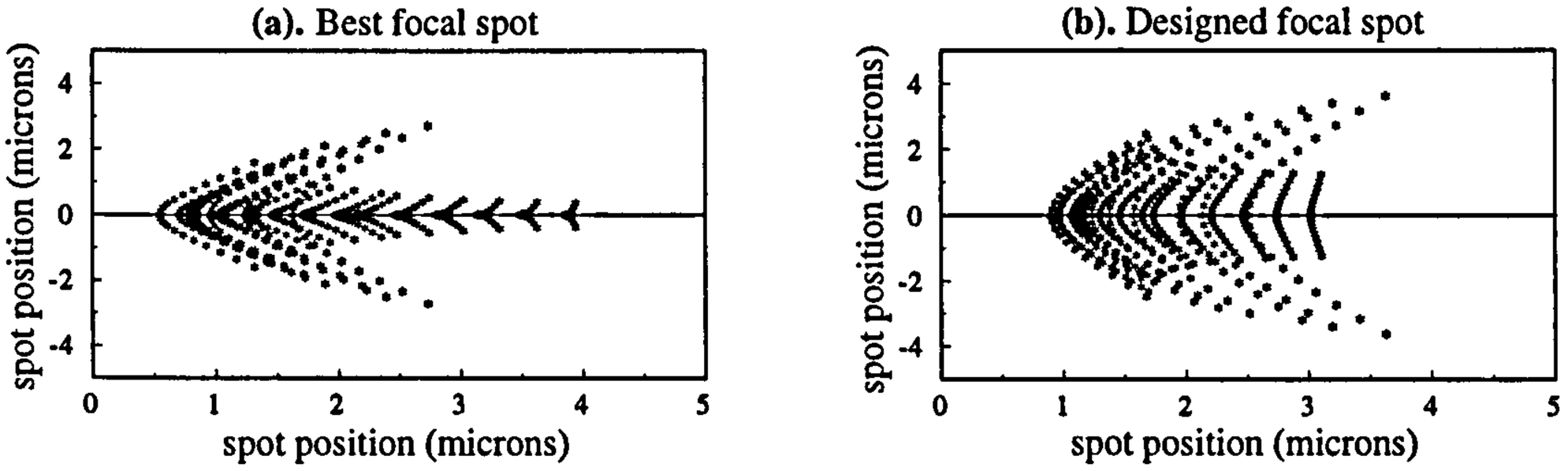


Figure 4.21: Replay spot diagrams traced from the recording points of figure 4.19: (a). Best focal spot, (from a focal plane search); and (b). Spot at the design focus.

The measured focal spot size, when illuminated with a uniform collimated beam at 850nm, was 8.7 $\pm$ 0.5μm (close to the convolved value of CodeV, see figure 4.22b), at a distance of 1210 $\pm$ 50μm from the gelatin surface of the hologram. This longer focal length (*cf.* 800μm design) is accounted for by the fact that the lens had to produce a focus through the hologram substrate, this increased the focal length by a factor of 0.5 over the thickness of the substrate, (1mm). Accounting for this refractive lengthening of the focal distance, the actual focus should be produced at a distance of 1137μm from the gelatin surface. The measured values of spot

size and focal length, allowing for the Rayleigh range (which for a uniform beam is $z_R = \frac{\omega_0^2}{0.744\lambda} = 60\mu$m) and experimental error, are very close to the theoretical predictions.

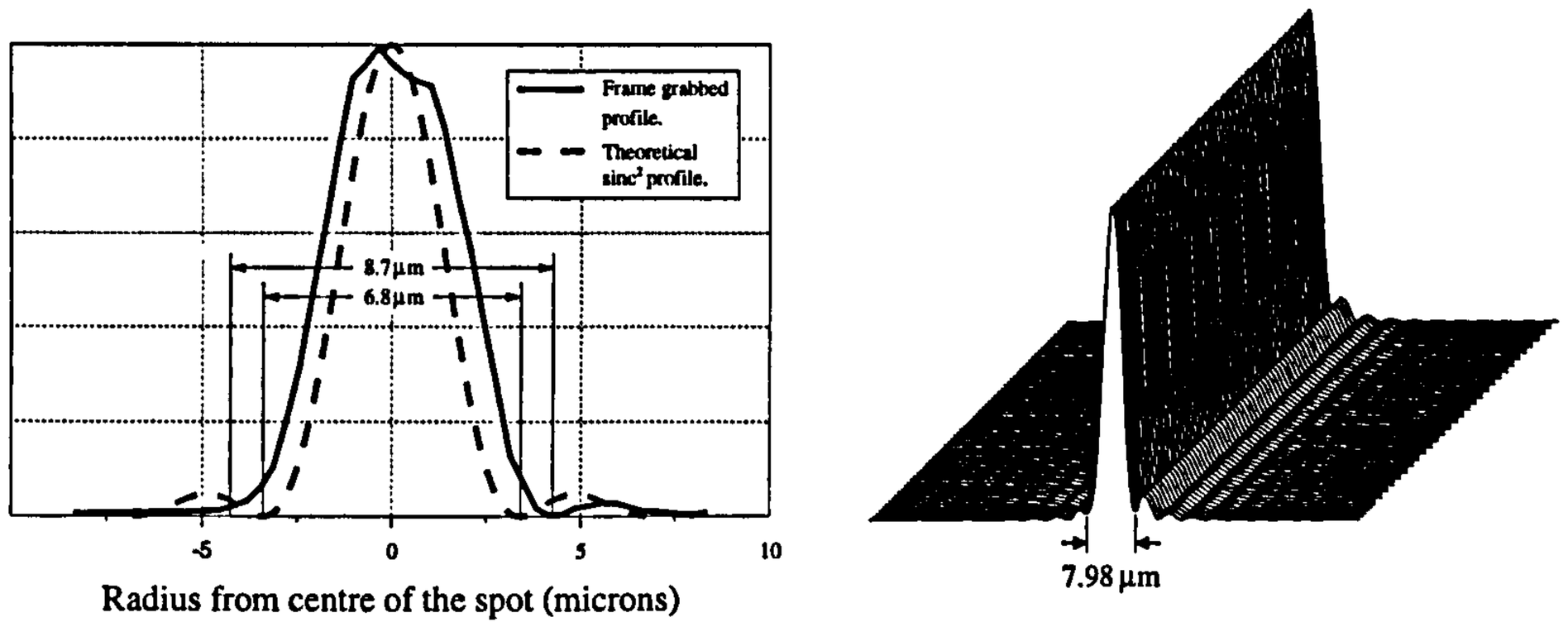


Figure 4.22: Spot intensity profiles: (a). Hamamatsu line scan through line focus vs. a normalised $sinc^2$ profile for the same aperture as the lens; and (b). ORA CodeV point spread function for the lens (incorporates aberrations with diffraction limit).

Figure 4.22a shows the line scan of a frame-grabbed image (using a Hamamatsu frame grabbing system) of the best focal spot together with an ideal normalised $sinc^2$ Fraunhofer pattern expected at this focal distance from a 200μm aperture. It is clear that the lens is not quite diffraction limited, lacking a secondary diffraction minimum to the left of the central peak. However, a $sinc^2$ function does not exactly apply to such diffractive lens elements, because of the off-axis nature of the incident beam and the asymmetric distances diffracted beams must travel before reaching the focus. This is evident from the close resemblance of the design focal spot diagram to the line scan (figures 4.21b & 4.22a). The experimental profile is broader because both profiles are normalised and the irregular nature of the peak appears to be due to the close-cascaded doublet structure. Although the lenslet is not perfectly diffraction-limited it does focus over 98% of the incident optical power into a spot 8.7μm in diameter.

The maximum Bragg mismatch was estimated (from the ray tracing) to be 0.51°, leading to a better than 99% theoretical diffraction efficiency overall. The efficiency of the lenslets was measured as 90%, with a uniformity of ±2%. The compensator

plate was recorded using the same arrangement as in figure 4.20, without the optics used to create the cylindrical wavefront and image the apertures. The aperture of the compensator plate is much greater than that of the entire lens array, consequently imaging was not required. The compensator plate had an efficiency of 95%, making the total efficiency of the cemented doublet 86%. The volume parameters of the DCG used to record this doublet HOE were found to be $\Delta n_{max} = 0.031$, (the difference between saturated refractive index and the base level), and $\beta = 1.2 \times 10^{-5} m^2/J$, (the emulsion sensitivity), resulting in a peak refractive index modulation of 0.027 for the maximum efficiencies achieved. Both elements were heat treated as per section 2.8 in order to protect them from thermal damage and to tune the Bragg angles with respect to one another.

There are two additional problems that can be encountered when cementing holographic elements together:

- **Moire fringes.** The periods of the gratings can beat in and out of phase with one another, creating regions of zero efficiency, unless they are carefully aligned. This can be overcome by aligning the uncured cemented doublets in the replay beam so as to eliminate all unwanted Moire effects and then UV curing the cement in situ. Moire fringes are usually visible under white light illumination but as can be seen from the photo of the cemented doublet, in figure 4.23, such fringes appear to have been avoided. However, the irregular nature of the line scan peak (figure 4.22a) may be due to a slight misalignment of the two HOEs.

- **Evanescent coupling.** Rayleigh orders from the compensator plate which would be evanescent when the HOEs are air-separated become coupled into the lens hologram due to the absence of the dielectric refractive index step. In addition, evanescent orders of the cylindrical back scattered waves in the lens become coupled into the compensator plate. This is a difficult situation to model mathematically and it will undoubtedly modify the replay behaviour of the lens under certain conditions. However, in the orientation in which this lens was designed to be used no unusual effects were apparent. The mea-

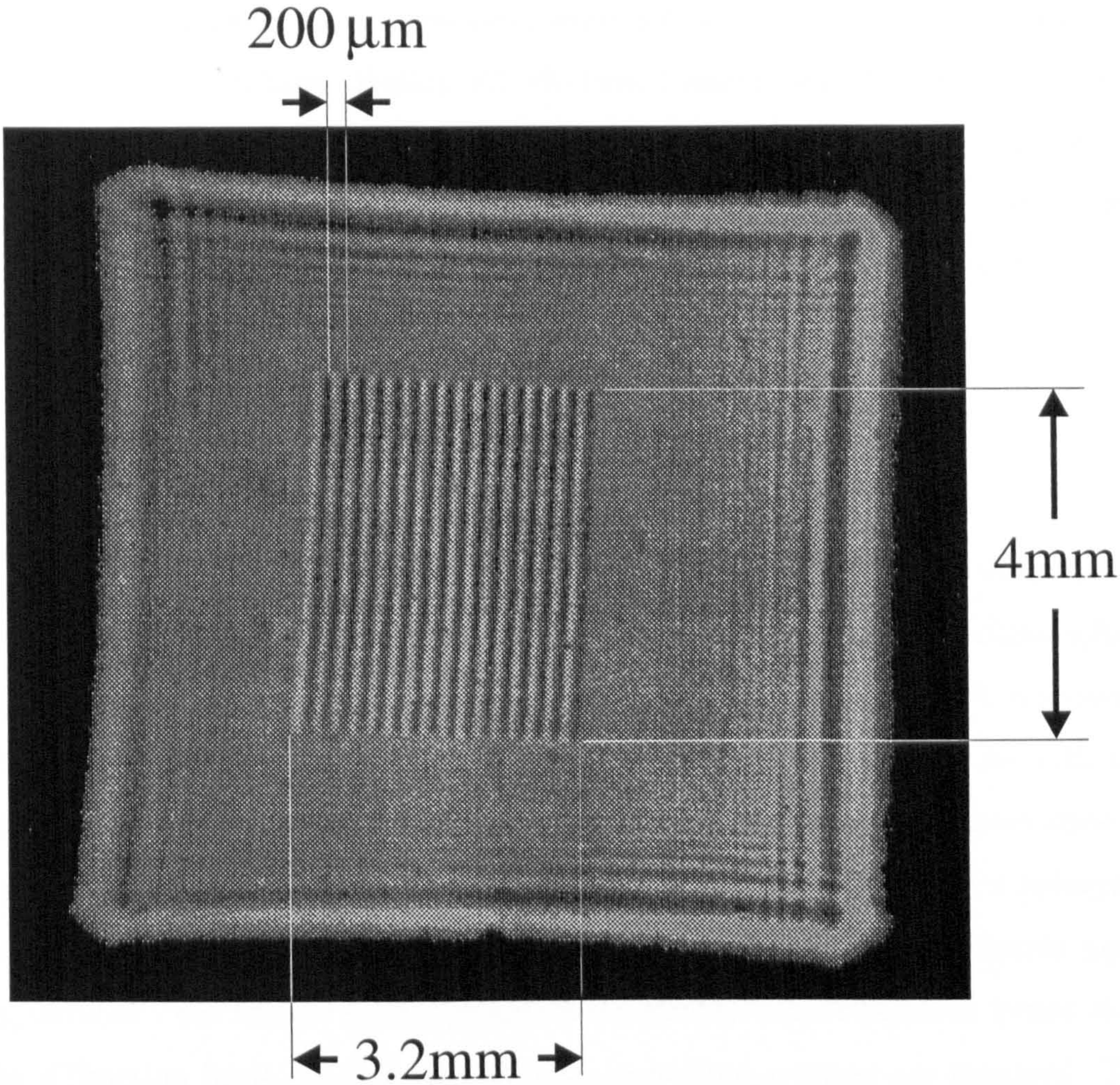


Figure 4.23: Cemented doublet of the 1 × 16 cylindrical lens array with on-axis compensator plate, illuminated with white light.

sured efficiency of the doublet was the same as that expected from the HOEs considered individually. Some work has been carried out on closely cascaded holograms [LAT72, STO86, SCH92] and it has been found that the geometric properties of the HOE are more tolerant to a change in replay wavelength, although the corresponding efficiency bandwidth is reduced. In addition, because this doublet is on-axis all beams travel more or less normally to the HOE surface resulting in the hologram being less polarisation dependent than off-axis lenses. Both of these factors are beneficial when dealing with laser diode sources; in particular making the HOEs in a communication system more tolerant to mode-hopping, (albeit at a reduced SNR).

To exactly quantify the aberrations produced by this lens an interferometric technique should be utilised. Hutley [HUT91] and Falkenstorfer [FAL91] respectively describe and quantify refractive and diffractive microlens aberrations with the use of both Twyman-Green and Mach-Zender interferometers. This lens worked well within its design constraints in the situation for which it was designed and so no further aberration analysis was carried out.

4.7 Summary

This new and much simplified method of wavelength-shifted focusing element design, produces results in good agreement with the other methods that utilise spherical (cylindrical) and planar wavefronts to record the HOE. Indeed, it represents the best results that are achievable using spherical recording optics. As with the other schemes this method is limited to producing only *microlenses* where close to diffraction-limited performance is required. When the physical size of a hologram becomes 'large', (the exact size at which an HOE becomes 'large' depends upon the f/number - see section 4.4.5), spherical wavefronts cannot produce lenses close to the diffraction limit. In this case aspheric recording systems are required. The theory in this chapter is laid out in a way that makes it easy to program as a design tool, (for both planar gratings and those with focusing power), and the close parallel to CGH elements is quite clear. Efficient on-axis lenses can be fabricated by the simple addition of a planar grating, Bragg matched to the input angle of the lens. This is easily recorded because, with this method, the recording angles required for the planar grating are the same as the average recording angle at the centre of the lens.

The method generates a simply computable recording geometry that can be used to fabricate efficient arrays of wavelength-shifted microlenses (with differing replay geometries - if required), exhibiting close to diffraction-limited performance. These can be used to produce multi-faceted interconnection arrays (see chapter 7) without resorting to the complex numerically intensive design algorithms and fabrication techniques of other methods.

Chapter 5

Fan-in and Fan-out

Most optical and electronic interconnection systems require some degree of fan-in and fan-out [GOO85]. The term fan-out refers to the splitting of a single interconnection channel into several channels each containing the same signal. Fan-in is the reverse of this, where separate signals are combined to form a single channel.

Optically, the fan-out of a single input beam of light to many outputs, such that the output beams replicate the input, can be achieved using diffractive components, waveguide junctions or (using bulk optics) multiple beam splitters. In principle fan-in can be accomplished by using fan-out systems in reverse. In addition, wavefronts can be split (or combined) into different power ratios or beam weightings, so that the output waves (for fan-out), although containing the same signal, do not have to be of equal optical power.

Although intuitively it would appear that optical signals can be fanned-in and out quite arbitrarily, this is not necessarily the case. The speed of electrical circuits is limited by the degree of fan-in/fan-out, (because of the impedance and capacitance defining a time constant in connecting channels), and a similar restriction applies to optical systems. This restriction is imposed by *the constant brightness theorem* which states that N beams can be fanned-in to the same spatial mode with only $1/N^{th}$ of the total incident power, i.e. no passive linear optical system can increase beam radiance, (a similar condition applies to fan-out). A discussion of its applicability to optical interconnects is presented in appendix C.

5.1 Fan-in

It is clear, from appendix C, that the constant brightness theorem is far more restrictive to fan-in than fan-out. The only way to completely avoid fan-in problems is to limit the fan-in level to two and use beams of orthogonal polarisation, (see chapter 7). Fan-in restrictions can, however, be circumvented in certain circumstances, using a variety of techniques [PRI88]. For a fan-in level of three or more there are trade-offs to consider associated with the switching speed of opto-electronic [MCC90] or opto-thermal [SMI87] devices that constitute the optical logic planes in computational or communications systems, (see the model in section 1.2). Larger device signal windows imply that the optical power is spread out over a larger area (reduced irradiance) and as a result the devices switch slower. One way around the fan-in limitations of the brightness theorem, for N beams, is to use a detector or device with an area N times greater than the individual beam sizes, (this often limits the bandwidth of the system). For example, if multiple beams are focused onto a device and the spot size of each beam is smaller than the device area then they may be directed to different parts of the device. Alternatively, multiple beams can be incident upon the same area of the device provided they occupy separate portions of the available solid angle. The resultant interference fringes average out over the area of the device, (e.g. the nearest-neighbour interconnect described in chapter 6 and optical matrix-matrix crossbar switch [BAR95, chapter 5]).

Additional problems can be encountered when fanning-in mutually coherent beams (e.g. when arrays of beams are generated from a single source using a binary phase grating) onto a diffractive element to realise a particular optical interconnect. Exact control of the individual beam phases must be maintained in order to avoid the appearance of interference fringes in the required output beams and the coupling of power into unwanted diffracted directions of the grating structure [LEG87, GLY89]. Such problems are not encountered with uncorrelated coherent signals [HIN94].

The rest of this chapter deals with an analysis of diffractive fan-out elements, (that can also be used to fan-in). In particular a model for volume holographic elements that demonstrates the possibility of high efficiency fan-out holograms with

preserved beam weightings.

5.2 Diffractive fan-out elements

There are many types of hologram that can be used to produce highly efficient optical fan-out in one or two dimensions [STR89], the most commonly used are the following:

1. Binary Phase Gratings (BPGs); these encompass Dammann [DAM71, TUR88] and Trapezoidal [MIL93, chapter 2] gratings. They are computer generated and usually fabricated using lithographic techniques [MIL93], though they can also be realised in DCG [ROB93a, chapters 5 & 6]. Diffraction efficiencies $\sim 65-75\%$ are achievable for 2D fan-out and $\sim 80\%$ for the 1D case [VAS92].

2. Multi-level (and continuous) phase gratings; commonly called *kinoforms* [LES69, VAS91]. They are designed and fabricated in a similar way to BPGs [MIL93, chapter 3] but offer a higher overall diffraction efficiency $\sim 90\%$.

3. Hybrid kinoforms; so called because they are fabricated using a hybrid of computer generated and conventional optical recording techniques [ROB91b, ICH92]. They can have a useful diffraction efficiency similar to volume HOEs $\sim 96\%$.

4. Modulated high-frequency carrier gratings; can be designed using three different coding systems - pulse-position [BLA95], pulse-width [STO91] and pulse-frequency [NOP94a, NOP94b]. They can be binary or multi-level phase gratings with efficiencies in the range $\sim 80-100\%$. They are computer generated and usually require direct electron or laser beam writing to fabricate them.

5. Multiplexed volume gratings; covered in the rest of this chapter.

The computer generated techniques of 1, 2 & 4 above have proved to be the most useful (particularly when generating a large 2D array of beams from a single source) not only because of the highly controlled fabrication technique that leads

to reproducible beam weightings (to within $\sim 5\%$), but also because CGHs are not restricted by the need to optically generate a wavefront in order to be able to record it. The disadvantage of these techniques when applied to optical interconnects is that all of the holograms are Fourier plane devices which cannot be used in some spatially variant interconnection schemes (see chapter 7).

5.3 Multiplexed volume fan-out gratings

Several planar volume gratings, (similar to those discussed in chapter 3), can be recorded in the same holographic emulsion. These gratings can each have their own separate object and reference recording beams or they can have separate object and a common reference beam. Intuitively the latter would appear to be a simple way in which to produce fan-out, since upon replay the multiple object beams will be reconstructed. There are, however, three key problems with this approach:

- Useful diffraction efficiency (the fraction of the input power coupled to the required output beams) is usually limited,
- The relative weightings of power in the various object beams are not generally preserved; and
- Spurious diffracted beams (not present at recording) can be generated, which if unwanted, constitute noise.

The sources of these problems have been addressed by many authors (see section 5.4.1), and will be discussed in sections 5.4 to 5.7 with respect to the model presented in this chapter.

The types of fan-out hologram generally of interest to digital optical computing or communication systems usually have a common reference beam and object beams centred symmetrically about the $z-$axis, (see figure 5.1). It is this type of fan-out HOE geometry which is primarily considered in this chapter.

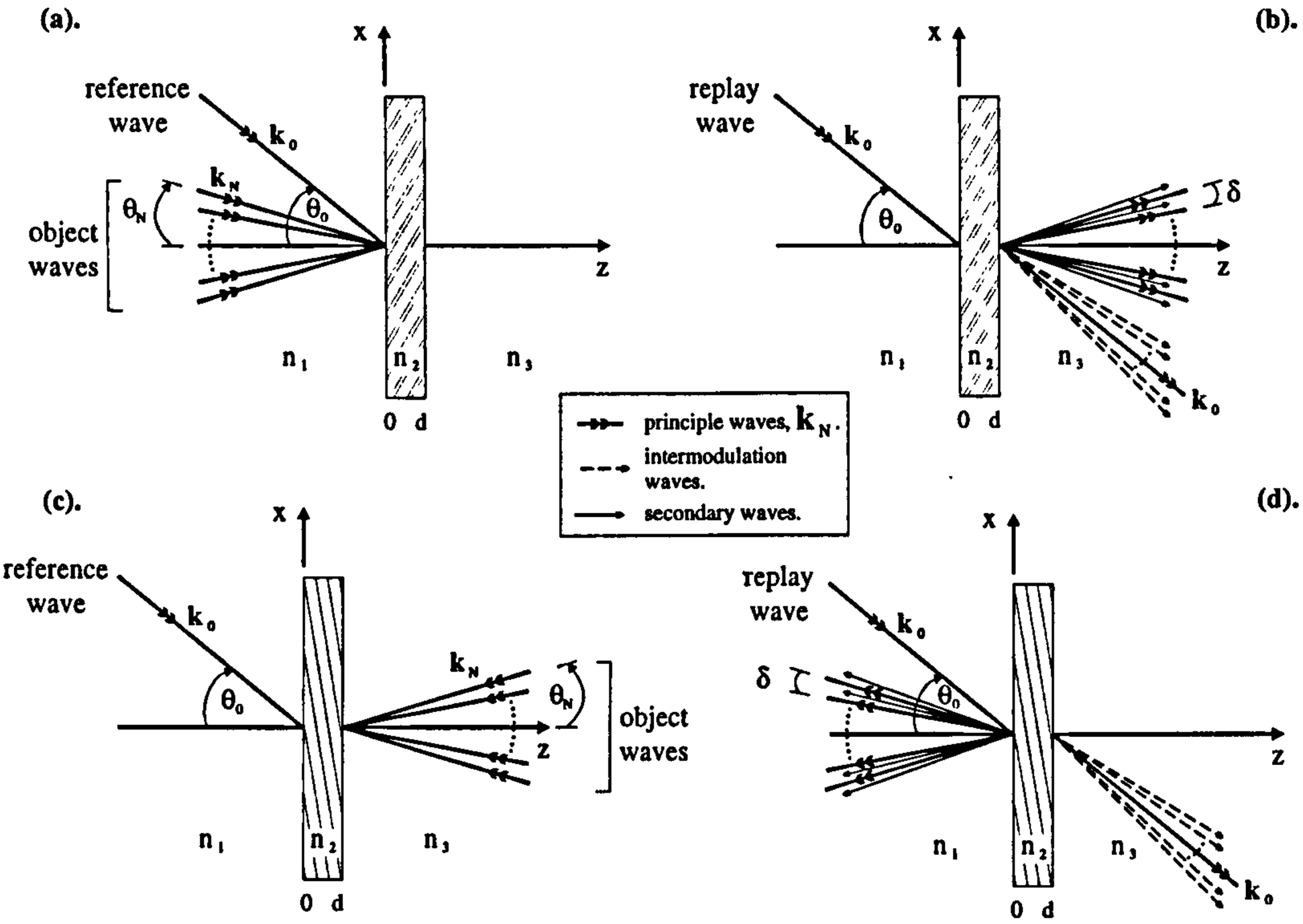


Figure 5.1: Schematic recording geometry of: (a). Transmission fan-out; and (c). Reflection fan-out gratings. The replay geometry and resulting principle and spurious waves are shown in (b). and (d). for transmission and reflection holograms respectively, ($\mathbf{k}_0$ is the reference or replay wavevector at an angle θ_0, $\mathbf{k}_N$ are the object wavevectors at angles θ_N and δ is the constant inter-object beam angle).

5.3.1 Recording multiplexed gratings

A multiple planar grating fan-out HOE can be recorded by the interference of a collimated reference wave, $E_0(\mathbf{r}) = E_0 \exp[-i\mathbf{k}_0.\mathbf{r}]$, and N object waves, $E_m(\mathbf{r}) = E_m \exp[-i\mathbf{k}_m.\mathbf{r}]$. This recording can take place either:

- **Sequentially,** which requires multiple exposures of the reference beam and each of the object beams in turn. There is no coherent relationship between the object waves, consequently this is often called *incoherent* recording, (only N primary gratings are formed). The electric field in the holographic plane, $E_{seq}(\mathbf{r})$, is defined as

$$E_{seq}(\mathbf{r}) = \sum_{m=1}^{N} (E_0(\mathbf{r}) + E_m(\mathbf{r})) \tag{5.1}$$

- **Simultaneously,** which requires a single exposure of the reference beam and all the object beams present at the same time. The object beams are then free to coherently interfere with one another and generate an additional $N(N-1)/2$ secondary gratings. Hence this is often called *coherent* recording. The electric field in the holographic plane, $E_{sim}(\mathbf{r})$, is defined as

$$E_{sim}(\mathbf{r}) = E_0(\mathbf{r}) + \sum_{m=1}^{N} E_m(\mathbf{r}) \tag{5.2}$$

The simultaneous recording technique will generate spurious diffracted beams because of the extra unwanted secondary gratings. It would appear that sequential recording should avoid such problems, but it has been pointed out by several authors, e.g. [RED89a], that this is not necessarily the case. The situation is not so simple and to understand why, a mathematical model of grating interactions is required, (see section 5.4.2).

The analysis contained in this chapter is largely theoretical and so a detailed explanation of recording techniques is not presented. However, the basic interferometric recording geometry, (similar to that illustrated in figure 2.6), can be extended to generate the multiple object beams, shown in figures 5.1a & 5.1c, needed to fabricate a fan-out hologram. Multiple sources can be generated using lenslet arrays [RED89a, page 126] or existing fan-out elements, e.g. Dammann gratings [HER92, ROB93a]. Sequential recordings can be made by masking the sources appropriately. Alternatively, tilting mirror [RED89a, page 129] or translatable lens (chapters 6 & 7) systems can be utilised in a sequential recording system to achieve the required θ_N object wave angles.

The desired beam weightings of a fan-out HOE upon replay are determined largely by the beam ratios at recording. These define the relative strengths of the various gratings and the resultant coupling between them.

5.3.2 Beam ratio

Initially discussed in section 2.9.3, with reference to simple planar gratings, the beam ratio is even more important to the discussion of fan-out elements. It is defined for simultaneous recording (with reference to section 5.3.1) as the ratio of power in the reference wave to that in the composite object beam:

$$B = \frac{|E_0|^2}{\sum_{m=1}^{N} |E_m|^2} \tag{5.3}$$

Kostuk [KOS89a] shows how this irradiance beam ratio corresponds to an electric field amplitude beam ratio and hence to a grating refractive index modulation ratio. For N equally intense object waves, creating N primary gratings with an index modulation of Δn_p and $N(N-1)/2$ secondary gratings with an index modulation of Δn_s, then the grating strengths are related to the beam ratio, B, by

$$\frac{\Delta n_p}{\Delta n_s} = \sqrt{NB} \tag{5.4}$$

Consequently, for sequential recording, where $\Delta n_s = 0$, the composite beam ratio, B, is equal to infinity, but the beam ratio for each individual exposure is usually unity so that all primary gratings are of equal strength, Δn_p.

Increasing the composite beam ratio, B, suppresses secondary gratings (sometimes called intermodulation or cross gratings) and consequently reduces light coupled into spurious diffracted beams. It can, therefore, be used to control the uniformity and efficiency of fan-out elements (see section 5.3.3). However, there is a trade-off to consider here, that will be discussed further in section 5.7.3: a higher beam ratio implies an increase in the background DC level of refractive index modulation. For a medium of limited dynamic range, such as DCG, this restricts the achievable diffraction efficiency as the emulsion saturates.

5.3.3 Efficiency and uniformity

Most fan-out holograms used in digital optical systems require uniform beam weightings upon replay; all beams contain the same optical power. Therefore, the replay quality of a fan-out element is measured not only by how efficiently it can couple power from the input into the output beams, but also by how uniformly the power is distributed amongst these beams.

It is useful to define two quantities in order to discuss the performance of fan-out elements:

- **Diffraction Efficiency,** η, which is most usefully defined as a "real" efficiency, (taking into account all losses at the element), by the ratio of power in the signal orders to the incident power:

$$\eta = \frac{\sum_{m=1}^{N} P_m}{P_T} , \tag{5.5}$$

  where P_m is the power of the m:th diffracted signal order and P_T is the total incident power.

- **Uniformity Error,** ΔU, sometimes called reconstruction error, which is defined as

$$\Delta U = \frac{P_{max} - P_{min}}{P_{max} + P_{min}} , \tag{5.6}$$

  where P_{max} is the maximum diffracted signal beam power and P_{min} is the minimum.

5.4 A model for volume HOE fan-out

In order to look at the behaviour of optical fan-out elements more closely a general mathematical model for multiplexed holograms is desirable. Such a model, incorporating *reflection*, as well as transmission geometries, and a fan-out level greater than 2, has not been published to date. This section addresses this need by further enhancing the rigorous coupled-wave method presented in chapter 3.

5.4.1 Volume fan-out HOE theories

Su and Gaylord [SU75], Case [CAS76] and Jaaskelainen & Kuittinen [JAA92] investigate the special case of diffraction from planar gratings with arbitrary profile that can be considered as a superposition of harmonic gratings. In such a situation the diffracted directions are fixed by the fundamental grating and cannot be chosen individually; this restricts the usefulness of such holograms in optical interconnection schemes.

Volume phase fan-out holograms, with several gratings multiplexed in the same volume have been studied by many authors. Table 5.1 summarises the key features of some of the main papers, (some study planar gratings, others utilise grating structures with focusing power), the trend being that later works tend to progress towards more accurate theoretical models. Where later papers appear to make no further advances over previous work, they contain other features, such as modelling recording medium non-linearities [BLA91].

The methods stated in the table relate to the volume diffraction theories discussed in section 3.1. However, since in this case there is more than one grating under consideration, the following distinction is made between CWMs when discussing multiple gratings: Kogelnik coupled-wave methods assume that only the waves present at recording will be reconstructed upon replay, whereas multi-wave CWMs allow for spurious waves (not present at recording) to exist - under certain circumstances these can have significant intensities. In addition to the work outlined in table 5.1, Slinger [SLI85] examined, in detail, various models used to accurately analyse multiple transmission phase gratings in silver halide emulsions.

The 3D vector RCWM of Glytsis & Gaylords [GLY89] is, in theory, able to model multiple superposed gratings and could be generalised to both coherent and incoherent recording. It is, however, so mathematically complicated and numerically intensive that it becomes impractical to analyse more than the two grating case.

This chapter develops a theory that attempts to transcend all of those shown in table 5.1. It is applicable to both transmission and reflection gratings, recorded both coherently and incoherently with a separate or common reference wave (although the

Author	Grating type: transmission, reflection *or* both	Recording method: incoherent, coherent *or* both	Reference beam: separate, common *or* both	Number of object waves	Experimental data	Theoretical Method
Case [CAS75]	transmission	incoherent	common	2	yes	Kogelnik CWM
Alferness & Case [ALF75a]	transmission	incoherent	separate	2	yes	Thin Grating Decomposition
Solymar [SOL77]	transmission	coherent	common	multiple	no	Kogelnik CWM
Kowarschik [KOW78a]	transmission	incoherent	separate	2	no	Kogelnik CWM
Kowarschik [KOW78b]	reflection	incoherent	common	2	no	Kogelnik CWM
Benlarbi & Solymar [BEN79]	transmission	coherent	common	2	no	Kogelnik CWM
Lewis & Solymar [LEW83]	transmission	incoherent	common	2	no	multi-wave CWM
Kostuk et al. [KOS86]	reflection	both	common	2	yes	multi-wave CWM
Slinger et al. [SLI86, SLI87]	transmission	both	common	multiple	yes	multi-wave CWM
Kostuk [KOS89a]	transmission	coherent	common	multiple	no	various†
Cawte [CAW91]	transmission	both	common	2	no	multi-wave CWM††
Blair & Solymar [BLA91]	transmission	incoherent	separate	2	yes	multi-wave CWM§
Redmond [RED89a]	transmission	both	common	multiple	yes	multi-wave CWM§§
Glytsis & Gaylord [GLY89]	both	incoherent	both	2♮	no	3D vector RCWM
Herzig et al. [HER92, HER93]	transmission	both	common	multiple	yes	multi-wave CWM‡

Table 5.1: Comparison of various volume phase fan-out hologram models, († covers 3 separate coupled wave models, †† derives analytical expressions for incoherent recording diffraction efficiencies that work for small fan-out angles, $< 1°$, § models a non-linear medium that produces real spurious gratings, §§ specific emphasis on fan-out for optical interconnection and inclusion of grating vector and bulk index non-uniformities, ‡ includes grating phases to optimise replay fidelity, & ♮ generalised to multiple object waves, but only applied to the two grating case).

analysis concentrates on a common reference) and multiple object waves using a 2D TE-mode rigorous coupled-wave theory, (developed from chapter 3). It is, however, only a 2D theory and is consequently limited to analysing fan-out in one dimension

only. Fan-out in 2D will produce broadly similar results [RED89a, page 134] and general trends can be inferred from this model.

5.4.2 RCWM generalised for multiple gratings

From the RCWT equation 3.12 it is possible to consider multiple absorption and/or permittivity modulated gratings in the same volume. The multiplexed holograms considered in this chapter are all sinusoidal phase gratings (without harmonic components) and consequently equation 3.13 can be simplified to

$$D_h = E_h = \Delta n_2(k_{02} - i\alpha_2), \tag{5.7}$$

and the expressions for permittivity and absorption (from equation 3.3), for each grating, simplify to

$$\begin{aligned} n &= n_2 + \Delta n_2 \cos(\mathbf{K}.\mathbf{r}) \text{ ; and} \\ \alpha &= \alpha_2 \end{aligned} \tag{5.8}$$

The full forms of these equations are equally valid in this generalisation, but the phase grating approach makes the presentation of the method simpler.

The strict formulation of the coupling equation presented in chapter 3 is relaxed. The wave-index, m, no longer represents the $2N+1$ diffracted wave orders (or space-harmonics, N either side of the zero-order) of the truncated numerical solution. Instead the wave-index represents the diffracted orders retained in the analysis from the general multiple interaction diffraction from a collection of arbitrary gratings. Under such conditions the RCWT equation 3.12 can be re-written as

$$\begin{aligned} \frac{d^2S_m(u)}{du^2} = \frac{2\xi_{2m}}{\kappa}\frac{dS_m(u)}{du} + \frac{(k_{02}^2 - \xi_{2m}^2 - \beta_m^2 - 2ik_{02}\alpha_2)}{\kappa^2}S_m(u) \\ + \frac{2\pi(k_{02} - i\alpha_2)}{\kappa^2\lambda}\sum_{a,\,b}\Delta n_a S_b(u) \end{aligned} \tag{5.9}$$

The summation represents the collection of waves, with amplitudes inside the medium of $S_b(u)$, that are directly coupled to space-harmonic m. Therefore, what

needs to be considered is which $S_b(u)$'s couple to each $S_m(u)$ and via which gratings, (represented by Δn_a). A convenient idea is that of the coupling matrix, highlighted in section 3.4, in which a sub-matrix of the RCWM state-equation matrix, b_{rs}, represents the summation of equation 5.9.

5.4.3 Coupling matrices

Of those authors who have carried out a multi-wave type of analysis using a CWM, many have chosen to formulate the coupling equations into a matrix form e.g. [LEW83, RED89a, CAW91, HER92]. In a RCWM analysis only a part of matrix b_{rs} is equated to a coupling matrix: that which represents the coupling between space-harmonic m and other space-harmonics. It can be seen that this is the bottom left constituent diagonal sub-matrix of equation 3.29.

A coupling matrix, C, can be defined as

$$C = \begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1N} \\ c_{21} & c_{22} & \cdots & c_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ c_{N1} & c_{N2} & \cdots & c_{NN} \end{bmatrix} \tag{5.10}$$

where

$$c_{mb} = \begin{cases} \chi_a, & \text{if wave } m \text{ couples to wave } b \text{ via grating } a \\ \zeta_m, & \text{if } m = b \\ 0, & \text{if there is no direct coupling} \end{cases} \tag{5.11}$$

and from equation 5.9,

$$\begin{aligned} \chi_a &= \frac{2\pi(k_{02} - i\alpha_2)}{\kappa^2 \lambda} \Delta n_a \,, \\ \zeta_m &= \frac{k_{02}^2 - \xi_{2m}^2 - \beta_m^2 - 2ik_{02}\alpha_2}{\kappa^2} \end{aligned} \tag{5.12}$$

By defining $\gamma_m = 2\xi_{2m}/\kappa$, equation 5.9 can be re-written as

$$\frac{d^2 S_m(u)}{du^2} = \gamma_m \frac{dS_m(u)}{du} + \zeta_m S_m(u) + \sum_{a,\,b} \chi_a S_b(u) \tag{5.13}$$

This in turn leads to a new formulation of the expanded state matrix equation 3.29, which now looks like this

$$\begin{bmatrix} \dot{S}_{1,0} \\ \dot{S}_{1,1} \\ \dot{S}_{1,2} \\ \dot{S}_{1,3} \\ \dot{S}_{1,4} \\ \vdots \\ \dot{S}_{2,0} \\ \dot{S}_{2,1} \\ \dot{S}_{2,2} \\ \dot{S}_{2,3} \\ \dot{S}_{2,4} \\ \vdots \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & & 1 & 0 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & & 0 & 1 & 0 & 0 & 0 & \\ 0 & 0 & 0 & 0 & 0 & \cdots & 0 & 0 & 1 & 0 & 0 & \cdots \\ 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 1 & 0 & \\ 0 & 0 & 0 & 0 & 0 & & 0 & 0 & 0 & 0 & 1 & \\ & & \vdots & & & & & & \vdots & & & \\ \lceil & & & & \rceil & & \gamma_0 & 0 & 0 & 0 & 0 & \\ & \ddots & & \ddots & & & 0 & \gamma_1 & 0 & 0 & 0 & \\ & \text{Coupling} & & & & & 0 & 0 & \gamma_2 & 0 & 0 & \cdots \\ & \text{matrix} & & & & & 0 & 0 & 0 & \gamma_3 & 0 & \\ & \ddots & & \ddots & & & 0 & 0 & 0 & 0 & \gamma_4 & \\ \lfloor & & & & \rfloor & & & & \vdots & & & \end{bmatrix} \begin{bmatrix} S_{1,0} \\ S_{1,1} \\ S_{1,2} \\ S_{1,3} \\ S_{1,4} \\ \vdots \\ S_{2,0} \\ S_{2,1} \\ S_{2,2} \\ S_{2,3} \\ S_{2,4} \\ \vdots \end{bmatrix} \tag{5.14}$$

Instead of drawing up the entire state-matrix when discussing wave-coupling it is convenient to simply refer to the coupling matrix.

The definition of a coupling matrix is simplified using the RCWM for TE polarisation because of a mathematical property specific to this theory. This key property (that does not apply to the TM theory) is that each equation for space-harmonic m only contains terms in the neighbouring diffracted orders. For example, the equation representing wave amplitude S_m only contains terms in S_{m-1} and S_{m+1}, (where harmonic gratings are not included).

The relaxing of the formulation used in chapter 3 simply implies that instead of $-N \leq m \leq +N$, now $0 \leq m \leq 2N$ for all space-harmonics considered in the analysis. Consequently, the $2N + 1$ space-harmonics can be any diffracted orders,

not just the first $\pm N$ (either side of the zero) generated in the more constrained formalism. This flexibility allows for a rigorous or loose analysis, as required, at the expense of the extra care needed to ensure that all grating couplings are correct. In addition, it allows troublesome evanescent orders that 'blow-up', (see section 3.5), and make the algorithm unstable, to be neglected (at the expense of rigour).

The two formalisms are equivalent for single gratings. Retaining all diffracted orders up to and including second-order space-harmonics in the analysis the two coupling matrices can be depicted in a tabular form:

m	-2	-1	0	1	2
-2	ζ_{-2}	χ_1	0	0	0
-1	χ_1	ζ_{-1}	χ_1	0	0
0	0	χ_1	ζ_0	χ_1	0
1	0	0	χ_1	ζ_1	χ_1
2	0	0	0	χ_1	ζ_{-2}

(a). RCWM of chapter 3

m	0	1	2	3	4
0	ζ_0	χ_1	χ_1	0	0
1	χ_1	ζ_1	0	χ_1	0
2	χ_1	0	ζ_2	0	χ_1
3	0	χ_1	0	ζ_3	0
4	0	0	χ_1	0	ζ_4

(b). Relaxed formulation

Table 5.2: Tabular form of the coupling-matrix, {a). the wave-vector representations from chapter 3; and b). the relaxed conditions of this chapter}, for space-harmonics upto and including the second diffracted order from a single grating, (where m is the space harmonic index).

Where the mapping between the two nomenclatures for space harmonics, m, is $0 \leftrightarrow 0, 1 \leftrightarrow -1, 2 \leftrightarrow +1, 3 \leftrightarrow -2$ and $4 \leftrightarrow +2$.

Harmonic gratings can be retained in the analysis (as described by Redmond [RED89a, pages 72-73]) with the same rigour as chapter 3. However, when multiple gratings are considered, diffracted directions become more complex than the simple space-harmonics generated by a single grating and the coupling matrix becomes a lot more complicated.

5.4.4 Diffracted directions

The vector Floquet condition, equation 3.8, becomes more complicated for multiple gratings, with the wavevector for each order being defined by which gratings it interacts with:

$$\mathbf{k}_m = \mathbf{k}_b + g\mathbf{K}_a \ , \tag{5.15}$$

where k_m is the resultant wavevector of a space-harmonic with wavevector k_b being diffracted by g orders of grating K_a. This leads to the 2D scalar propagation constants, of section 3.2.3, being redefined as

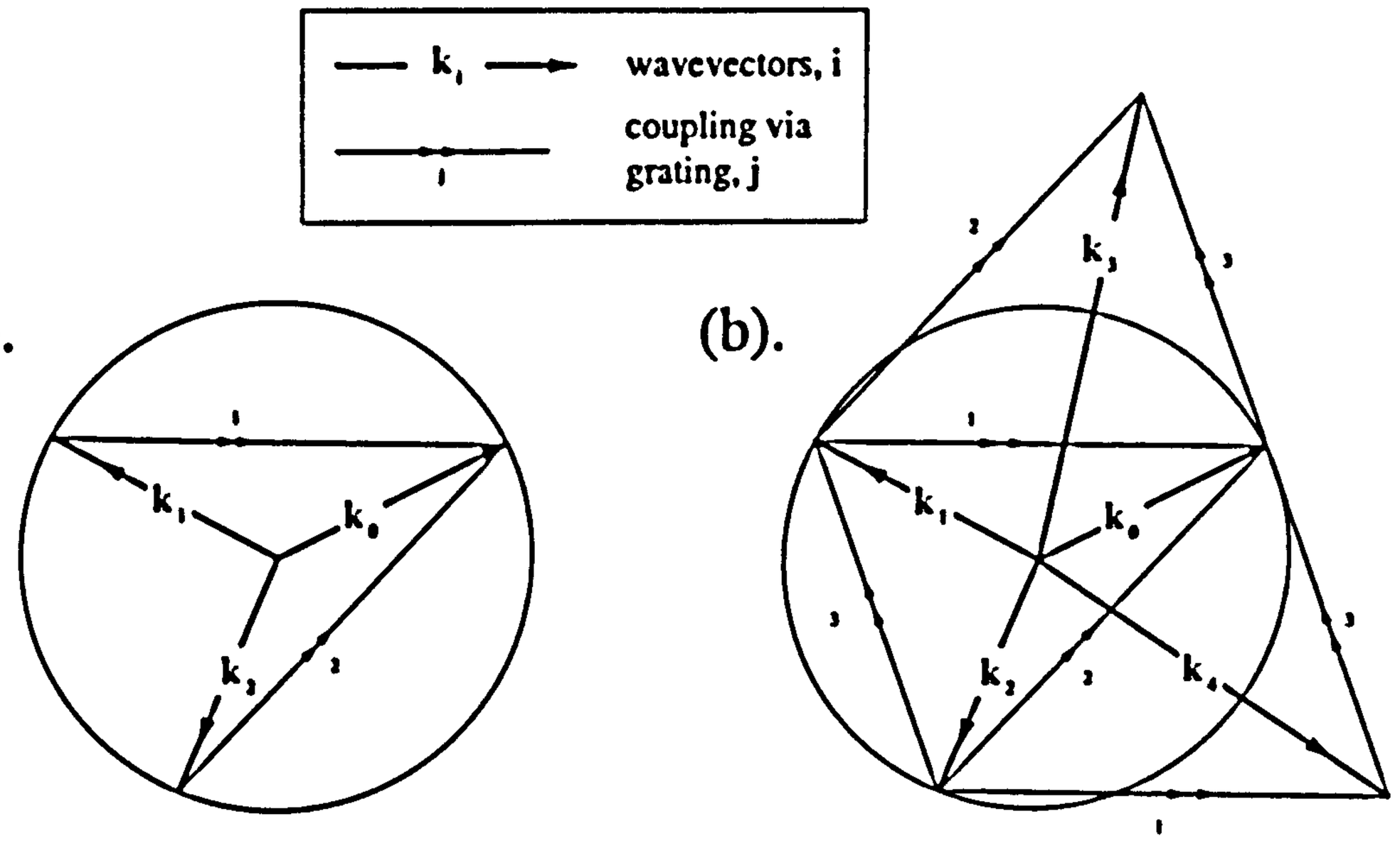
$$\begin{aligned} \xi_{2m} &= k_{bx} + gK_a \cos\phi_a \text{ ; and} \\ \beta_m &= k_{bx} + gK_a \sin\phi_a \text{ ,} \end{aligned} \tag{5.16}$$

for all space-harmonics, m. The $z-$component of the wavevectors inside and outside of the modulated region, ξ_{1m} and ξ_{3m}, are the same as before. A strict definition of a wavevector for waves, of the same frequency, propagating through a homogeneous dielectric medium, requires the length of such vectors to be the same when depicted in wavevector diagrams. Vector Floquet diagrams, (see figures 5.2 & 5.3) illustrate diffracted space-harmonic vector directions (as defined by the vector Floquet condition) for a modulated dielectric medium. Such diagrams are often referred to as wavevector diagrams and the space-harmonic vectors termed wavevectors. From figure 5.2b it is clear that all such "wavevectors" are not the same length and therefore cannot strictly be called wavevectors. However, the convention of naming these wave directions *wavevectors* (rather than Floquet or space-harmonic vectors) is used in this chapter.

As pointed out by Glytsis & Gaylord [GLY89], the total number of possible space-harmonics inside the medium increases from ∞ to ∞^2 with the simple addition of a second grating in the same volume. For more than two gratings the problem scales geometrically (N gratings generating ∞^N orders), consequently the truncation of the analysis must be carefully chosen to retain accuracy but still remain computable.

Under the formulation, presented here, the wave index, m, becomes an arbitrary assignment and careful note must be made of which order corresponds to which level of diffraction and from what grating. The method of Glytsis & Gaylord avoids this by generating an additional wavevector subscript for each diffraction from a new grating. This, however, becomes infeasible as the number of gratings increases and they had difficulty applying their method to more than two gratings.

The coupling matrix, allowed diffracted directions, desired recording waves and



$$k_1 = k_0 - K_1 \qquad k_3 = k_0 - K_1 + K_2 = k_0 + K_3$$
$$k_2 = k_0 - K_2 \qquad k_4 = k_0 - K_2 + K_1 = k_0 - K_3$$

Figure 5.2: Wavevector diagram of the reconstruction of: (a). the recording waves by the replay wave; and (b). the recording and first-order intermodulation waves by the replay wave, in a reflection fan-out to two hologram.

the strength of the gratings can be manually manipulated into a generalised RCWM program or an algorithm can be used to generate the matrix and space-harmonics automatically.

To arrive at a methodology for defining diffracted directions from a multiple grating hologram, let us first examine the simplest incoherent two-grating case. Consider a two grating reflection hologram, the vector Floquet diagram of which is depicted in figure 5.2a. From the revised vector Floquet condition, equation 5.15, the replay beam, k_0, generates the two recording beams k_1 and k_2. It has been pointed out by several authors, e.g. [SLI86, RED89a], that each of the newly generated waves, k_1 and k_2, may interact with the grating corresponding to the other, provided the Bragg responses of the two gratings overlap. This 'double diffraction' gives rise to two new waves, arbitrarily assigned wavevectors k_3 and k_4. These vectors have

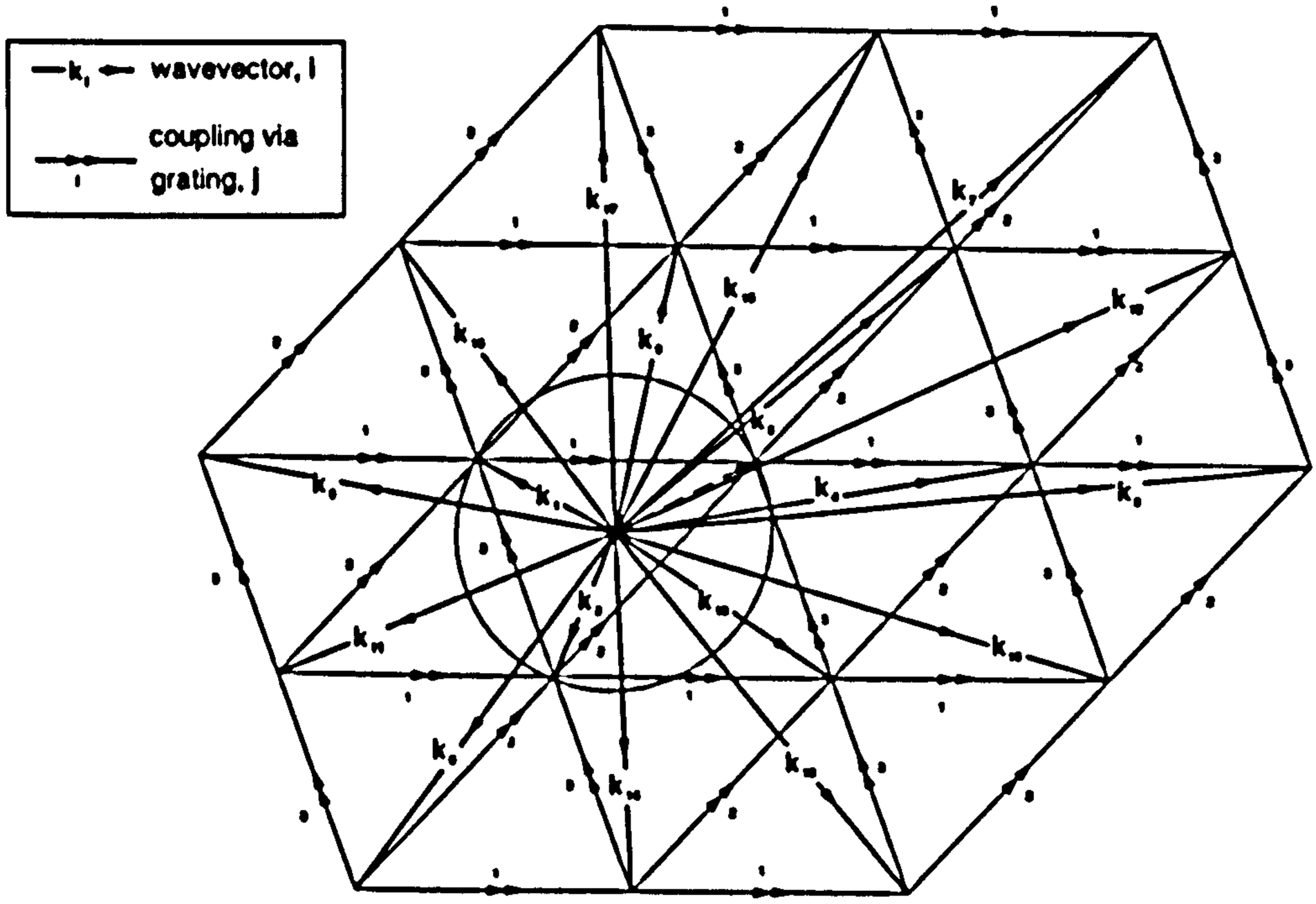


Figure 5.3: Vector Floquet diagram for reflection fan-out to two, including every order generated upto and including second-order diffraction by all gratings (grating coupling is illustrated by a double-arrow with the grating number next to it and the zero-order is shown as a dotted line).

directions given by the vector Floquet condition as

$$\begin{aligned} k_3 &= k_1 + K_2 = k_0 - K_1 + K_2 \text{ ; and} \\ k_4 &= k_2 + K_1 = k_0 - K_2 + K_1, \end{aligned} \tag{5.17}$$

as shown in figure 5.2b.

If instead the hologram is recorded coherently, a third grating K_3 is generated between k_1 and k_2, as is also included in figure 5.2 b. It can be seen from simple rules of vector addition and equation 5.15 that the extra diffracted orders, k_3 and k_4, can now equally be defined as

$$\begin{aligned} k_3 &= k_0 + K_3 \text{ , and} \\ k_4 &= k_0 - K_3 \text{ ; since} \\ K_3 &= K_1 - K_2 \end{aligned} \tag{5.18}$$

Therefore the *same* waves can be generated in both the incoherent and coherent recording situations and an ambiguity exists (from equations 5.17 and 5.18) in the way that the newly diffracted directions can be described.

k	0	1	2	3	4	5	6	7	8	۹	10	11	12	13	14	15	16	17	18
0	ζ_0	χ_1	χ_2	χ_2	χ_1	0	0	0	0	χ_3	χ_3	0	0	0	0	0	0	0	0
1	χ_1	ζ_1	χ_3	0	0	χ_1	0	0	0	χ_2	0	χ_2	0	χ_3	0	0	0	0	0
2	χ_2	χ_3	ζ_2	0	0	0	χ_2	0	0	0	χ_1	χ_1	0	0	χ_3	0	0	0	0
3	χ_2	0	0	ζ_3	χ_3	0	0	χ_2	0	χ_1	0	0	χ_1	0	0	χ_3	0	0	0
4	χ_1	0	0	χ_3	ζ_4	0	0	0	χ_1	0	χ_2	0	χ_2	0	0	0	χ_3	0	0
5	0	χ_1	0	0	0	ζ_5	0	0	0	0	0	χ_3	0	χ_2	0	0	0	0	0
6	0	0	χ_2	0	0	0	ζ_6	0	0	0	0	χ_3	0	0	χ_1	0	0	0	0
7	0	0	0	χ_2	0	0	0	ζ_7	0	0	0	0	χ_3	0	0	χ_1	0	0	0
8	0	0	0	0	χ_1	0	0	0	ζ_8	0	0	0	χ_3	0	0	0	χ_2	0	0
۹	χ_3	χ_2	0	χ_1	0	0	0	0	0	ζ_9	0	0	0	χ_1	0	χ_2	0	χ_3	0
10	χ_3	0	χ_1	0	χ_2	0	0	0	0	0	ζ_{10}	0	0	0	χ_2	0	χ_1	0	χ_3
11	0	χ_2	χ_1	0	0	χ_3	χ_3	0	0	0	0	ζ_{11}	0	0	0	0	0	0	0
12	0	0	0	χ_1	χ_2	0	0	χ_3	χ_3	0	0	0	ζ_{12}	0	0	0	0	0	0
13	0	χ_3	0	0	0	χ_2	0	0	0	χ_1	0	0	0	ζ_{13}	0	0	0	χ_2	0
14	0	0	χ_3	0	0	0	χ_1	0	0	0	χ_2	0	0	0	ζ_{14}	0	0	0	χ_1
15	0	0	0	χ_3	0	0	0	χ_1	0	χ_2	0	0	0	0	0	ζ_{15}	0	χ_1	0
16	0	0	0	0	χ_3	0	0	0	χ_2	0	χ_1	0	0	0	0	0	ζ_{16}	0	χ_2
17	0	0	0	0	0	0	0	0	0	χ_3	0	0	0	χ_2	0	χ_1	0	ζ_{17}	0
18	0	0	0	0	0	0	0	0	0	0	χ_3	0	0	0	χ_1	0	χ_2	0	ζ_{18}

Table 5.3: Tabular form of the coupling-matrix for space-harmonics, up to and including the second diffracted order, from a a reflection fan-out to two hologram, (index k represents the wavevectors of figure 5.3).

This double diffraction process can go on (according to a rigorous analysis) with each newly diffracted direction being diffracted from each of the two-primary or single-secondary gratings (either directly or indirectly) ad-infinitum. The optical power in the 'spurious' diffracted directions depends upon the strengths of the secondary and primary gratings and the degree of Bragg overlap. These factors will specify which space-harmonics need to be maintained in the analysis to achieve accuracy and convergence with the RCWM. Predictably, they are difficult to define. The fan-out to two reflection hologram analysis is depicted in figure 5.3 for all generated directions, up to and including second-order, from both the primary and secondary gratings. The coupling matrix describing this wavevector model can be represented in a tabular form as table 5.3. If the gratings are recorded incoherently then the

coupling constant for the secondary grating is set to zero, i.e. $\chi_3 = 0$.

Clearly the analysis becomes quite complicated for the two-grating case; for three gratings the problem is significantly worse since for N primary gratings the number of possible secondary gratings is $N(N-1)/2$.

If an accurate vector Floquet diagram can be drawn for a specific multiple grating situation, then space-harmonic and coupling information can be extracted from it and manually entered into a generalised RCWM program. However, for a degree of fan-out beyond three the analysis is usually too intensive to deal with manually. A general rule for the degree of diffracted directions necessary to take into account, and an algorithm to generate these directions and the respective couplings, is required.

5.4.5 Discriminatory wavevector coupling algorithm

From an analysis of two and three grating situations it became clear that for volume phase gratings replayed at or close to Bragg incidence the key diffracted directions and couplings that must remain in the analysis are defined by the following constraint:

$$\begin{aligned} &-1 \text{ order from the } \textit{primary} \text{ grating ; and} \\ &\pm 1 \text{ orders from the } \textit{secondary} \text{ grating} \end{aligned} \tag{5.19}$$

This level of wavevector and coupling truncation is in agreement with the work of Herzig et al. [HER92, HER93] who assume that primary gratings are generally optically thick and secondary gratings are mainly optically thin (see appendix A). A detailed example of the algorithm used in this chapter, under the truncation implied by constraint 5.19, is presented in appendix D.

The following clarifies the way wavevectors and coupling are described with reference to appendix D.

Wavevector directions

Unwanted (spurious) wavevectors fall into two broad categories:

- **I-waves.** Intermodulation waves generated by diffraction of the zero-order (or replay wave) either directly from a secondary grating or indirectly via two

primary gratings; and

- **S-waves.** Secondary waves generated by the double diffraction (either directly or indirectly) from space-harmonics other than the zero-order.

Thus -1 primary and ± 1 secondary grating order diffraction are used to generate all the possible diffraction orders. Some of these will be degenerate. For example, in the fan-out to two case (illustrated in figure 5.3) the -1 order secondary diffraction from $\mathbf{k}_1$ via grating $\mathbf{K}_3$ generates the same wavevector as $\mathbf{k}_2$. However, some ambiguities that can arise from the many ways of defining the same wavevector, (e.g. from figure 5.3: $\mathbf{k}_{13} = \mathbf{k}_2 + 2\mathbf{K}_3$ or $\mathbf{k}_{13} = \mathbf{k}_1 + \mathbf{K}_3$) are eliminated by constraint 5.19 because at most $g = \pm 1$.

Taking the reflection fan-out to two shown in figure 5.3 as an example, the wavevectors break down into the following categories:

- Primary waves: $\mathbf{k}_0$ the replay wave, $\mathbf{k}_1$ and $\mathbf{k}_2$ the recording waves, $\mathbf{k}_{3..8}$ the harmonics from the primary gratings,
- I-waves: $\mathbf{k}_9$, $\mathbf{k}_{10}$, $\mathbf{k}_{17}$ and $\mathbf{k}_{18}$; and
- S-waves: all others.

However, under constraint 5.19 only a subset of these wavevectors are retained by the algorithm, these are shown (re-labelled) in figure 5.8b.

The total number of space-harmonics, m_{max}, (coupled-wave equations) generated as a function of the level of fan-out (or primary gratings), N, is then

$$m_{max} = (N+1) + (N-1)N(N+1) \tag{5.20}$$

The wavevectors that are degenerate are then picked out and eliminated from the analysis, reducing m_{max} accordingly (see appendix D).

Coupling

It is a complicated problem to automatically generate all of the allowed couplings between wavevectors. The constraint of equation 5.19 makes this task easier but it is

still not simple. All of the information needed to generate the coupling is contained in tables such as that shown in figure D.1; these can be constructed quite simply for any level of fan-out. Coupling information is contained in the descriptions of degenerate orders and so these must be calculated and the information extracted before they can be disregarded.

From the work of Glytsis & Gaylord [GLY89] and figure 5.3 it is clear that the k-space representation of the multiple grating vectors is a regular geometric grid of inter-meshed grating vectors, defined by the rules of vector addition. It is possible using these rules and the vector Floquet condition (equation 5.15) to calculate a representation of the coupling between the diffraction orders retained in the analysis. This, however, does not represent the true coupling matrix. To obtain a full coupling description the degenerate orders must be eliminated and the coupling that exists between them and the other non-degenerate orders copied to each of the remaining representations of the degenerate orders held in the analysis. This results in a complete coupling matrix for the -1 primary and ± 1 secondary grating diffraction. For a detailed example of how the algorithm works see appendix D.

This approach to the RCWM is generally applicable and rigorous, though the algorithm truncates the number of wavevectors to a suitable level for volume fan-out. It can be made more accurate, particularly for optically thin gratings, by the inclusion of higher order diffracted waves - modifying constraint 5.19 and extending the discriminatory algorithm rules outlined in appendix D. The penalty paid for this is increasing the eigenvalue problem, perhaps beyond the memory and accuracy limits of many computers.

5.5 Numerical solution

The model was implemented in a similar way to that of chapter 3 (sections 3.4 & 3.4.1). However, the size of the eigenvalue problem is considerably larger than before and this size scales disproportionately with the level of fan-out. The eigenvalue problem requires the manipulation of large 2D double-precision (complex for the inclusion of absorption gratings) arrays and therefore a lot of virtual mem-

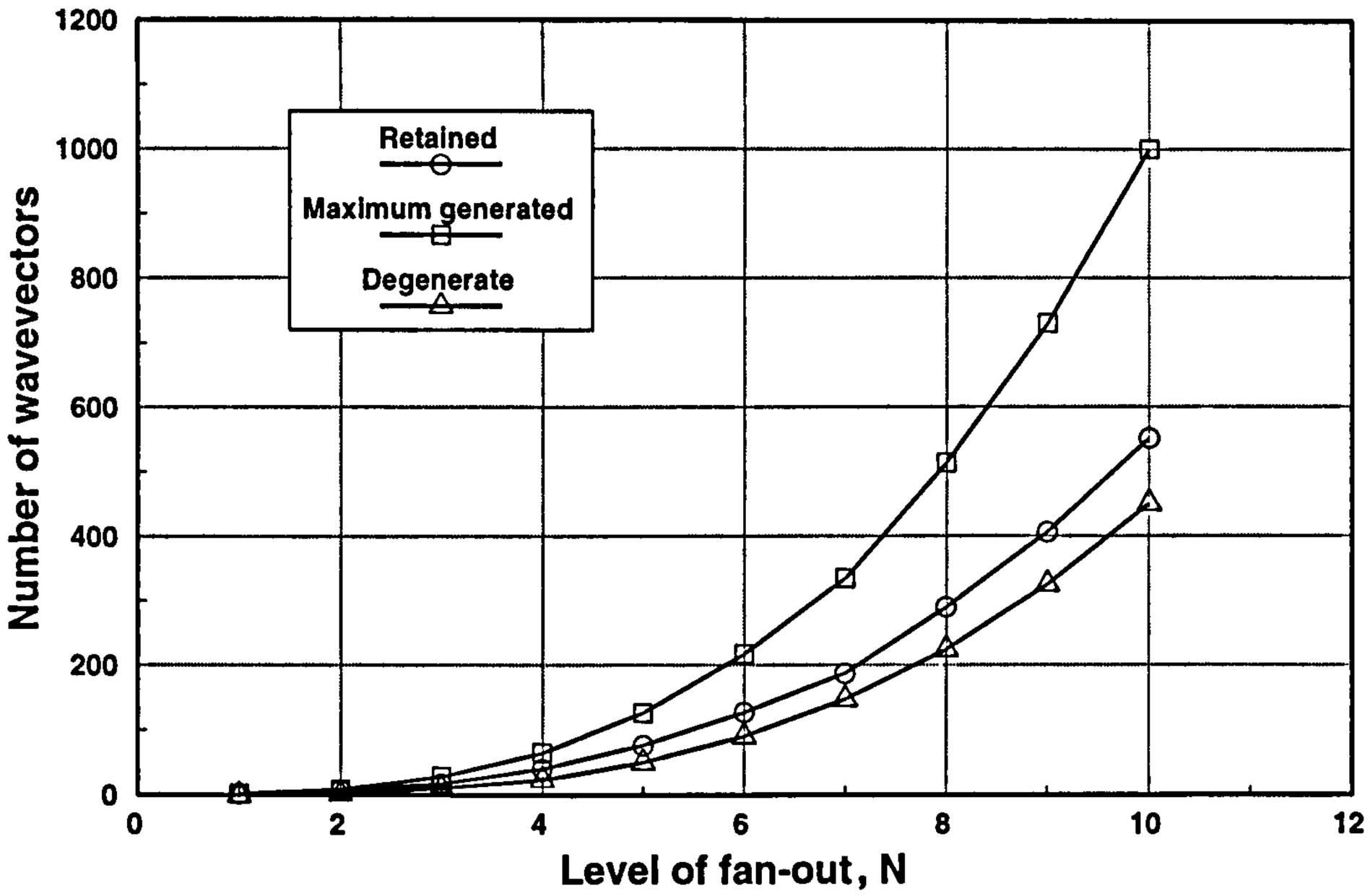


Figure 5.4: The maximum number of wavevectors generated, m_{max}, for a particular level of fan-out (N primary gratings), the number of which are degenerate and the resulting number retained in the analysis.

ory and, for the SUN system on which the model was run, a considerable amount of swap space. For example, if $N = 10$ then $m_{max} = 1001$ requiring a 2D array of 2002×2002 to store b_{rs} of equation 5.14 and with 8 bytes for each value, this is a total of nearly 32Mb of memory for just one instance of this array. Figure 5.4 shows how the number of wavevectors in the analysis varies with the level of fan-out for: the maximum number defined by equation 5.20, the degenerate waves derived from the discriminatory algorithm (see appendix D) and the resulting number of waves retained once the degeneracy has been removed. Predictably, it can take many hours of CPU time to carry out some of the analysis required for even moderate levels of fan-out.

5.6 Experimental verification

The following section tests the agreement between the theoretical results produced by this model and some of the experimental and theoretical fan-out data of other

authors.

5.6.1 Verification for transmission fan-out

Transmission fan-out holograms have been analysed in far more detail and by many more authors than their reflection counterparts (see table 5.1). A very detailed analysis of volume fan-out for optical interconnects has been carried out by Redmond [RED89a, chapter 4]. A duplication of Redmond's analysis has been carried out with this model and an excellent degree of agreement reached. The duplication is not presented here, but the minor discrepancies are highlighted. Section 5.7.2 contains a comparison with the work of Redmond for incoherent fan-out at small fan-out angles.

Redmond's model is a multi-wave CWM and contains many of the assumptions about grating orders and the degree of coupling made in the truncated analysis of this model, (sections 5.4.3 to 5.4.5), with one major difference: the wavevectors maintained in his analysis correspond to the primary and intermodulation waves only. Higher order intermodulation waves (produced by the $\pm$2nd order grating interactions) were also looked at, however, secondary waves were not included in Redmond's analysis.

Secondary waves are of particular importance when the multiple gratings are recorded simultaneously (coherently). In this case, the beam ratios between object waves are ~ 1, recording a strong secondary grating. These strong secondary gratings can couple a lot of power away from the primary (or principal) waves (into the secondary waves), reducing the useful diffraction efficiency and altering the uniformity of those waves. The effect is the same as that observed in reflection fan-out gratings with similar beam ratios, discussed in section 5.7.

The conclusions that can be drawn from this are:

- This model is in broad agreement with the experimental and theoretical observations of Redmond for incoherently recorded transmission volume fan-out elements.

- Secondary waves are important when fan-out holograms are recorded coher-

ently, (where inter-object beam ratios approach unity); and

- When analysing multiple transmission gratings a more simple multi-wave CWM would be equally applicable, provided secondary waves are included in the analysis. However, a multi-wave CWM is not appropriate for the *general* modelling of multiple reflection gratings because of the slowly varying amplitude approximation (see section 3.1.4 and appendix A). This RCWM, however, can be applied to both transmission and reflection fan-out holograms.

Physical and processing parameters.	
Recording angles	$\theta_R = -3°$, $\theta_1 = -13°$ & $\theta_2 = -26°$
refractive indices	$n_1 = n_3 = 1.0$ & $n_2 = 1.64$ (before and after processing)
absorption	$\alpha = 0.048\mu m^{-1}$
emulsion thickness	$d = 6.0\mu m$ (before processing) & $\Delta d = 0.085\mu m$ (1.42% after processing)
sequential modulation	$\Delta n_{21} = 0.0185$ & $\Delta n_{22} = 0.0170$
simultaneous modulation	$\Delta n_{21} = 0.0202$, $\Delta n_{22} = 0.0202$ & $\Delta n_{23} = 0.0230$

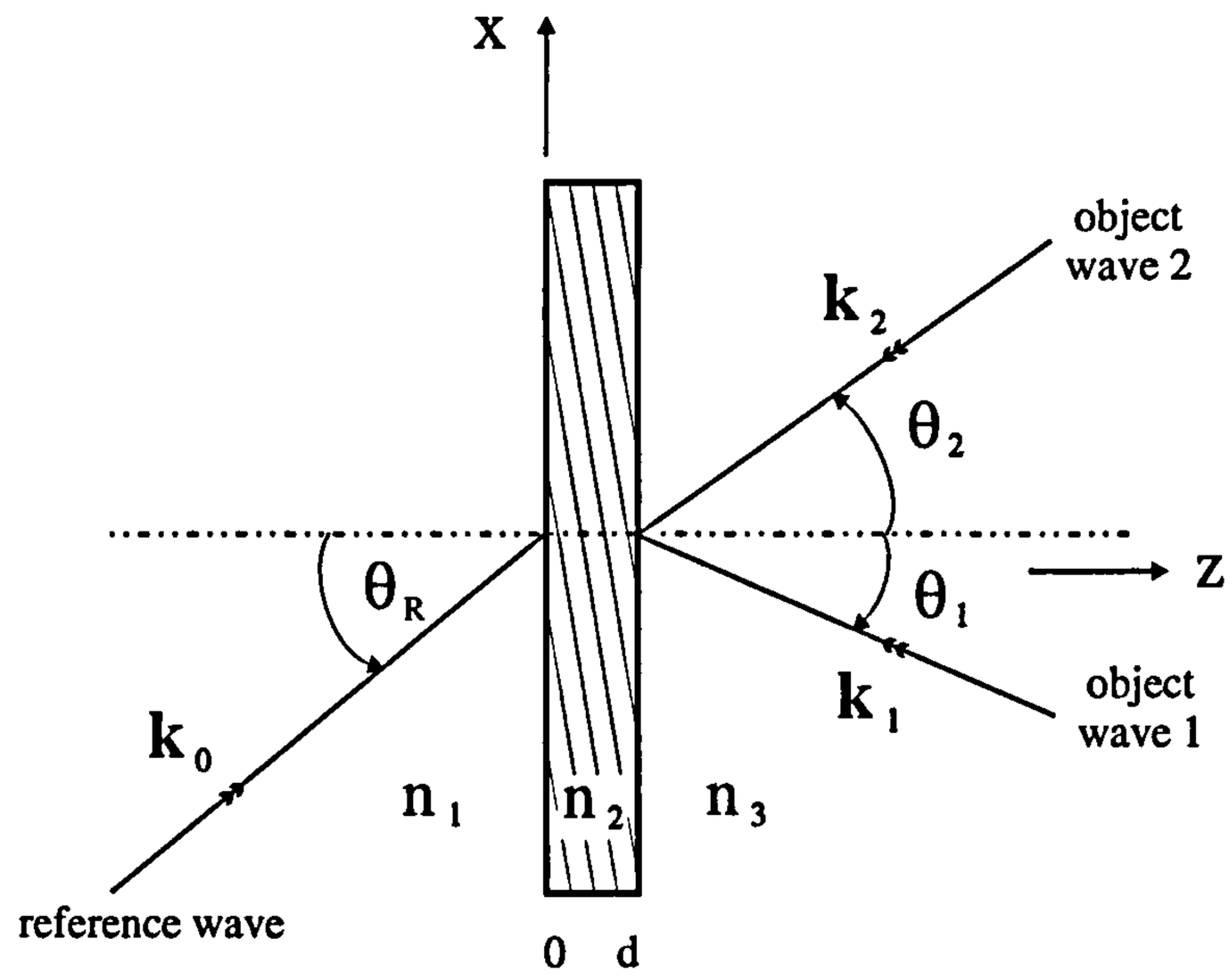


Figure 5.5: Multiple reflection grating recording arrangement of Kostuk et al. [KOS86], together with the physical and processing parameters (shown in the table insert [table 5.4]) used by them.

5.6.2 Verification for reflection fan-out

Only two authors have explicitly attempted to analyse multiple reflection gratings in the same volume: Kowarschik [KOW78b] and Kostuk et al. [KOS86]. Both restrict themselves to the two grating case and do not attempt to model higher degrees of fan-out. The theory of Glytsis & Gaylord [GLY89] concentrates on transmission gratings but could be applied to reflection gratings; however, it would have difficulty dealing with more than the two grating case. Kowarschik presents theoretical results for a simple coupled-wave model of sequentially recorded gratings only, whereas

Kostuk et al. utilise a multi-wave CWM and show experimental and theoretical data for both sequential and simultaneous recording. A direct comparison between the theoretical and experimental results of Kostuk et al. and the RCWM model of this chapter is presented here. The comparison is made to verify that this RCWM is in agreement with published reflection fan-out data, before a more in depth theoretical analysis of multiplexed reflection gratings is presented (see section 5.7).

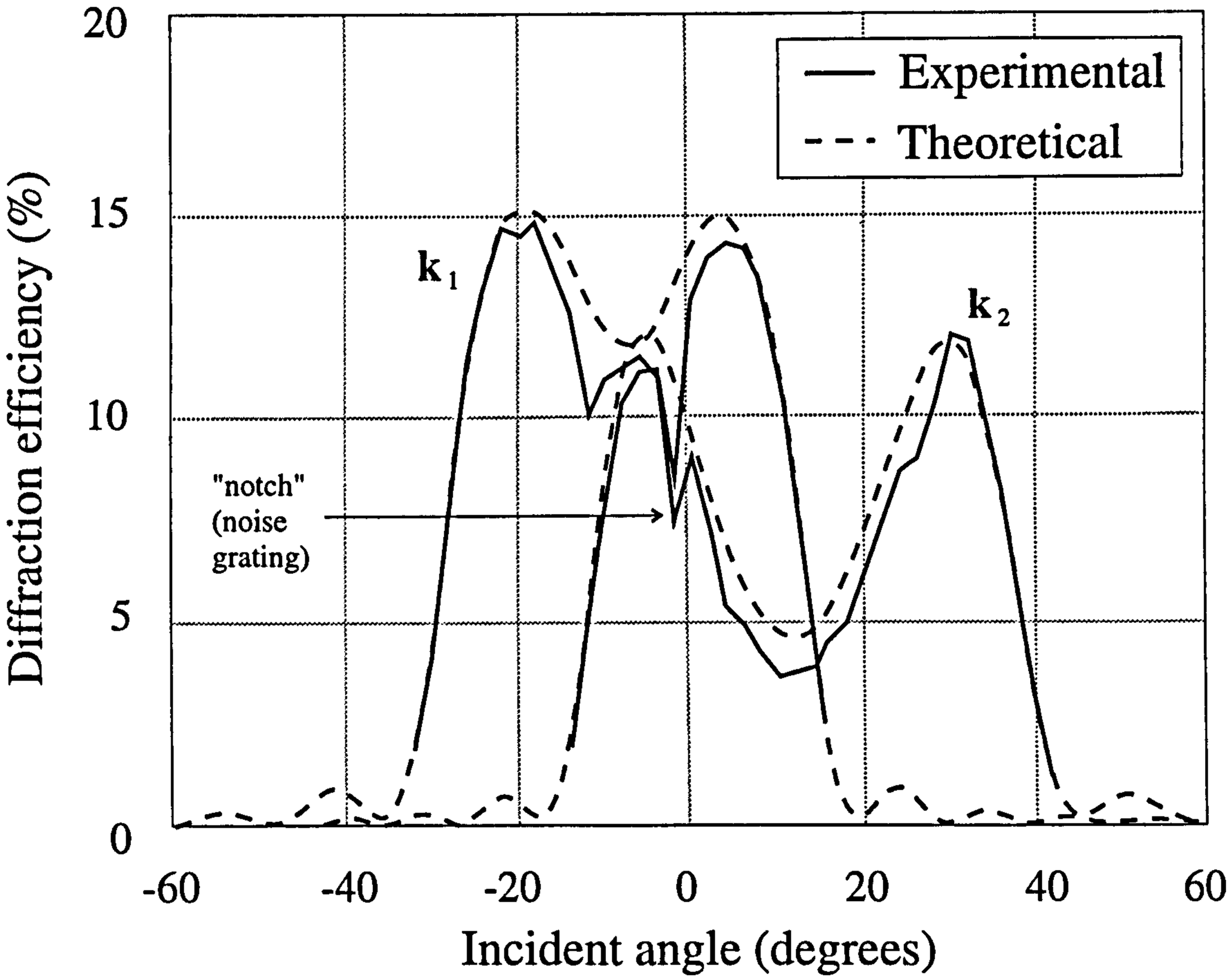


Figure 5.6: Experimental and theoretical results of Kostuk et al. [KOS86], for the *sequential* recording of a reflection fan-out to two hologram. The **k** vectors labelled above correspond to those shown in figure 5.8a and the incident angle corresponds to the replay wavefront (or reference plane wave) labelled $\mathbf{k}_0$ in figure 5.5.

Experiment and theory of Kostuk et al.

A fan-out to two reflection hologram is recorded (by Kostuk et al. [KOS86]) in a bleached silver halide emulsion (Agfa-Gevaert 8E75HD), by multiplexing the two gratings both simultaneously and sequentially. The recording geometry, along with

the physical and processing parameters of the emulsion are shown in figure 5.5. The authors omit the recording and replay wavelengths that they used, however, a good agreement with their results is achieved using the RCWM model assuming that both wavelengths are 633nm (He-Ne laser source). This seems feasible when considering the wavelength sensitivity of the emulsion that they used [BIE77, page 25].

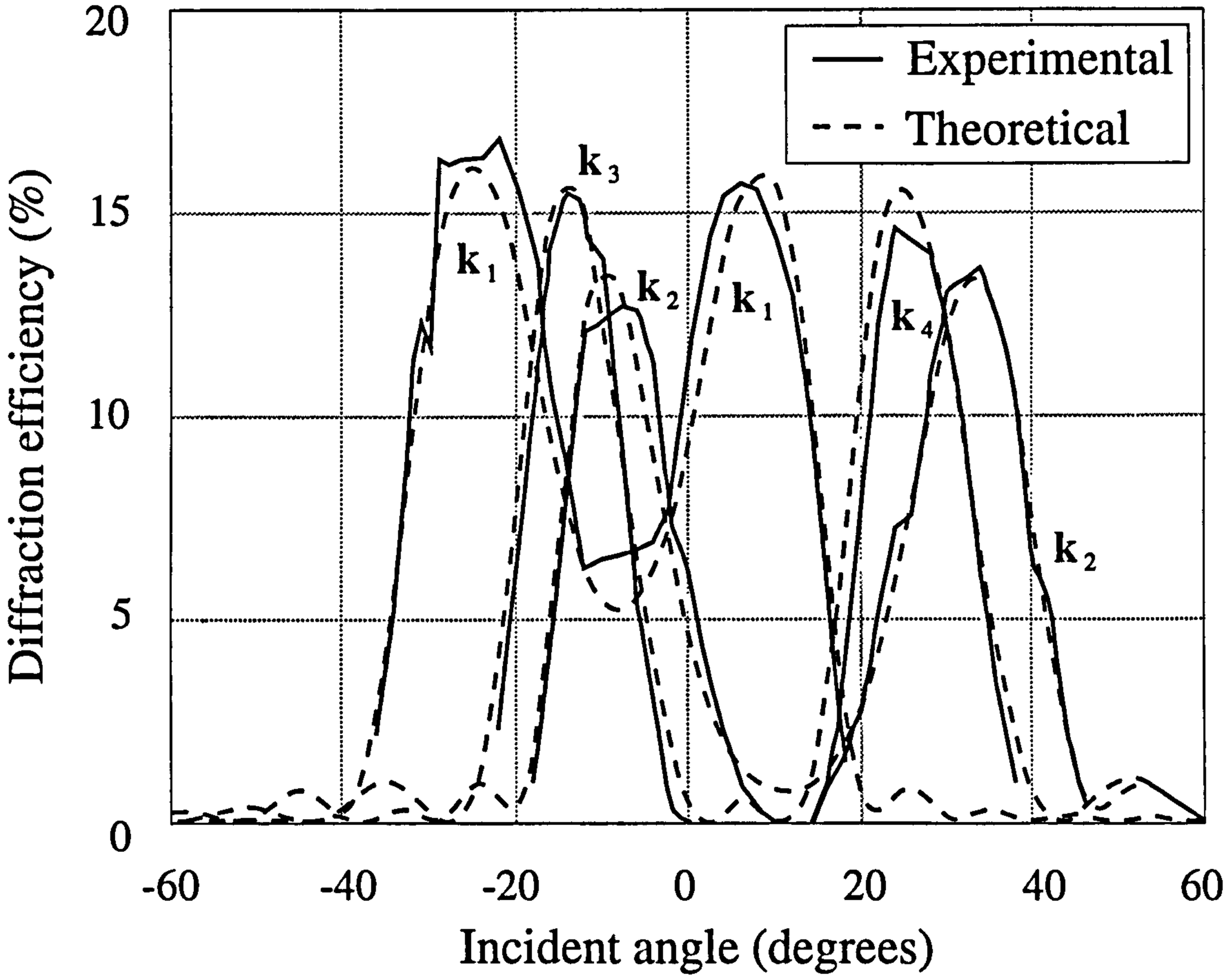


Figure 5.7: Experimental and theoretical results of Kostuk et al. [KOS86], for the *simultaneous* recording of a reflection fan-out to two hologram. The **k** vectors labelled above correspond to those shown in figure 5.8a and the incident angle corresponds to the replay wavefront (or reference plane wave) labelled $\mathbf{k}_0$ in figure 5.5).

For the main diffracted orders, over a range of reconstruction angles, there was a good agreement between their theoretical and experimental diffraction efficiencies. The angular spectra presented by Kostuk are shown in figures 5.6 and 5.7 for the incoherent and coherent recording situations respectively. Wavevectors $\mathbf{k}_1$ and $\mathbf{k}_2$ are the primary diffracted orders (reflected) and $\mathbf{k}_3$ and $\mathbf{k}_4$ are the first order intermodulation wavevectors (transmitted, either side of the zero-order or replay beam);

figure 5.8a shows a wavevector diagram of these space-harmonics.

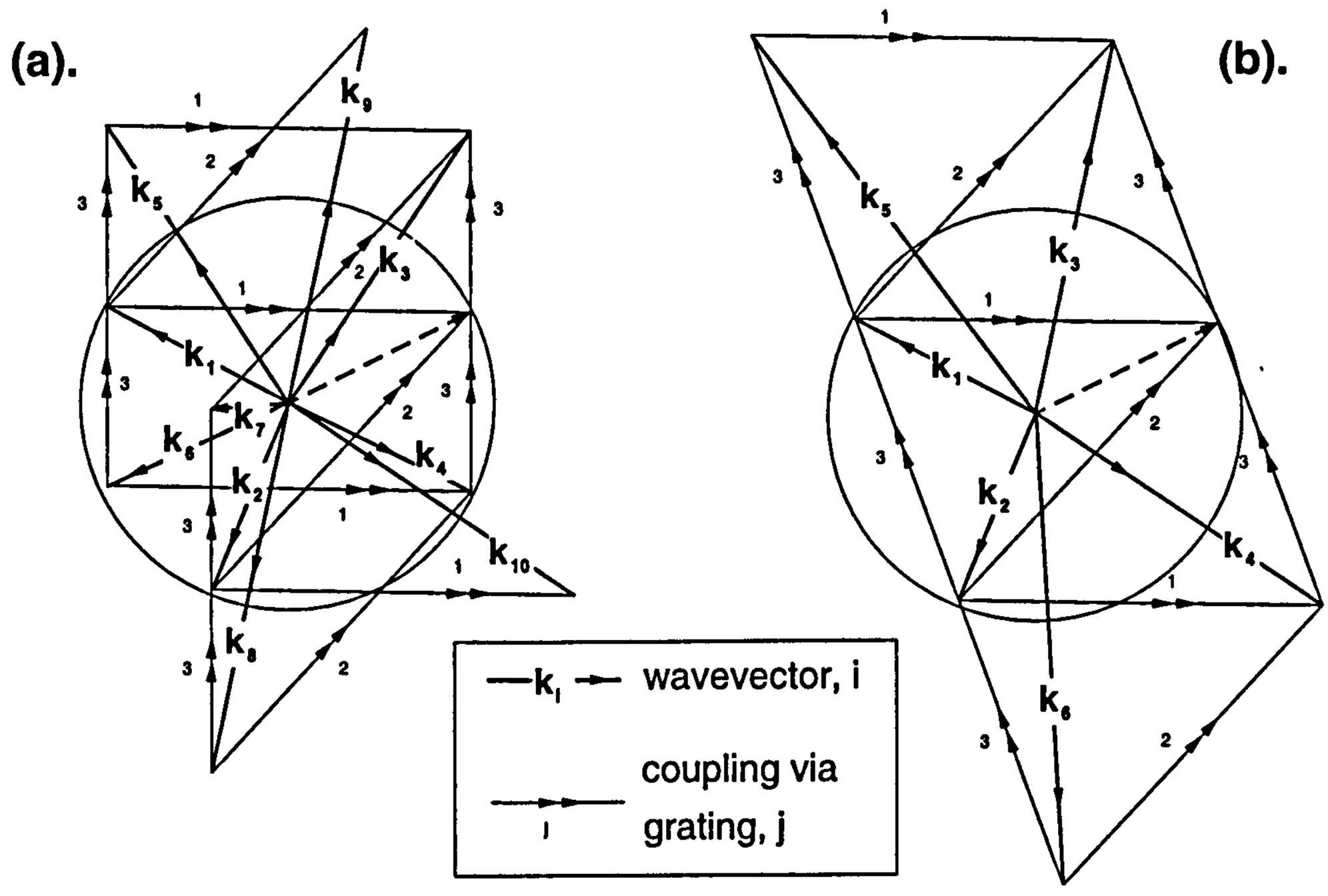


Figure 5.8: Wavevector diagrams showing the coupling of: (a). Kostuk et al. and (b). this chapter; included separately into the RCWM fan-out model to give the results shown in figures 5.9, 5.11 and 5.10, (the zero-order, replay beam, is shown dotted).

Coupling discrepancy

The theoretical model used by Kostuk does not use the space-harmonic coupling discussed in this chapter, (see figure 5.3 for the fan-out to two case). The coupling assumed by Kostuk is illustrated in figure 5.8a and (for the same level of space-harmonic generation) figure 5.8b shows the standard coupling used in this chapter. Clearly the simple triangle of grating vectors shown in figure 5.2b is not used by Kostuk et al. Their explanation of this is that the wavevector diagram of figure 5.8b is only applicable when the replay conditions exactly duplicate the recording conditions and since (in the experiment) an emulsion thickness change occurs, this cannot be the case. This assumption means that the degeneracy of wavevectors, discussed in section 5.4.5, no longer applies.

If an emulsion thickness (or bulk index) change is uniform, this assumption of

Kostuk cannot be true because such a change will alter all grating vectors by the same amount in the $z-$direction, thus maintaining the closed triangle of vectors shown in figure 5.2b. Consequently, the coupling would be as described in the rest of this chapter. However, if this coupling (and wavevector degeneracy) is used in the following comparison with the RCWM model for simultaneous recording, the theoretical calculations do not match the experimental data (see figure 5.10); the coupling does not matter for sequential recording since no real cross grating is generated. Therefore, there must be a slight thickness/fringe structure variation over the aperture of the HOE to bring about the coupling (required to match the experimental data), used by Kostuk. Another possible explanation could be the introduction of harmonic gratings since the silver halide has a total modulation $\Delta n \sim 0.06$, (see section 5.7.3). The grating vector mismatch will only be very slight (figure 5.8a greatly exaggerates the situation), but it is sufficient to break down the degenerate coupling structure of uniform gratings.

The coupling shown in figure 5.8 is therefore generally correct for uniform gratings. In this case, however, the coupling of figure 5.8a is required to match the experimental data.

RCWM method applied to the results of Kostuk et al.

The model presented in this chapter is applied to the results of Kostuk using the same parameters (as shown in figure 5.5) unless stated otherwise. The substrate of the HOE is modelled assuming that the dielectric region before the hologram is air and the region after has an index of 1.51, (the refractive index of Agfa 8E75HD substrates).

Taking the sequential and simultaneous situations separately:

1. **Incoherent recording:** The theoretical fit of the RCWM model with the experimental data is shown in figure 5.9; they are clearly in good agreement. The coupling discrepancy discussed above does not affect these results, it is only of consequence when a real cross-grating is recorded. The model uses slightly different physical parameters from those of Kostuk: instead of the val-

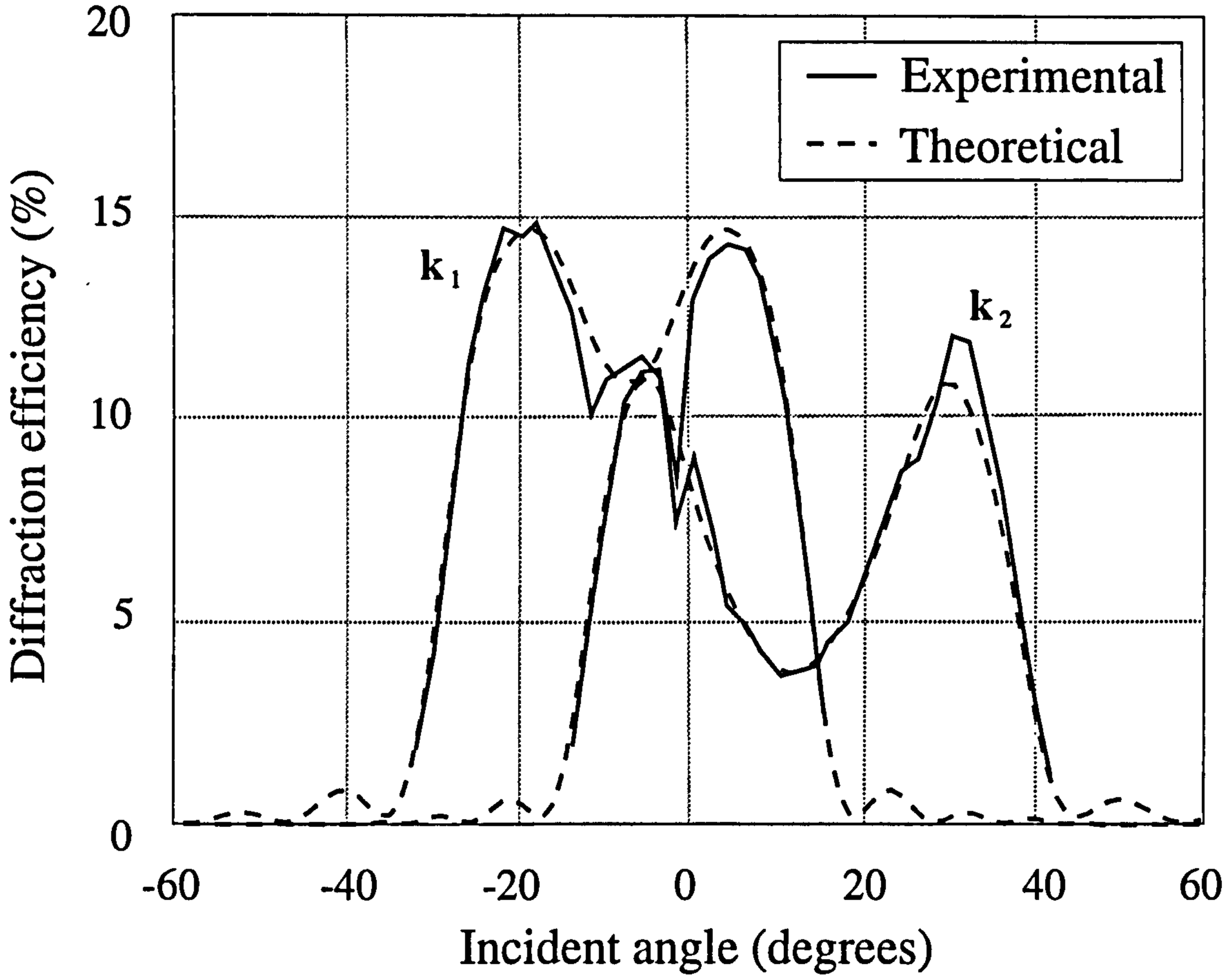


Figure 5.9: Theoretical RCWM comparison of results with the experimental data of Kostuk et al. [KOS86], for the *sequential* recording of a reflection fan-out to two hologram. The **k** vectors and coupling labelled above correspond to those shown in figure 5.8a.

ues shown in figure 5.5 the emulsion swelling is taken as 0.6% and the index modulation of primary gratings $\mathbf{K_1}$ & $\mathbf{K_2}$ are 0.0202 and 0.0175 respectively. The index modulations (used by Kostuk) were derived empirically by curve fitting, so it is unsurprising that the values are different from those used here, (the modulation imbalance between the gratings is explained in [KOS86] as the result of reciprocity failure in the silver halide). The thickness change discrepancy can be explained by the fact that the recording of a real transmission cross-grating adds mechanical rigidity to the structure in the emulsion. Consequently the incoherent recording must be more prone to a collapsing of the grating structure than the coherent recording, hence the reduced post-process swelling. The theoretical fit is at least as good as Kostuk, (allowing for the noise grating *notches* explained in reference [KOS86]).

2. **Coherent recording:** The theoretical results assuming the coupling and wavevectors of figure 5.8b are shown in figure 5.10; the agreement between experiment and theory is poor. Clearly the coupling assumed by Kostuk et al. is correct for their experimental HOE. Using the same coupling as Kostuk and the wavevectors shown in figure 5.8a the theoretical model gives the results shown in figure 5.11. Again the theoretical fit is at least as good as Kostuk and can be improved further if 25 wavevectors are included in the analysis. However, when this number of wavevectors are included, the evanescent wave instabilities, discussed in section 3.5, limit the scope of the theoretical angular spectrum. It is clear that, using the coupling of Kostuk, some secondary wavevectors are diffracted in directions close to the primary waves. In particular, from figure 5.8a, $\mathbf{k}_7$ will be very close to $\mathbf{k}_1$ (but still decoupled and distinct). The result of this could be that power in $\mathbf{k}_1$ is spatially indistinguishable from power in $\mathbf{k}_7$ and this could account for the "hump" in the experimental diffraction minimum of the angular spectrum of $\mathbf{k}_1$. Similar arguments can be made, for the small discrepancies between experiment and theory, with respect to other primary and secondary wavevectors.

The conclusions that can be drawn from this comparison with the work of Kostuk et al. are:

- This RCWM model shows an excellent degree of agreement between the experimental and theoretical results of [KOS86].

- The coupling used by Kostuk is of relevance to the particular experiment being analysed. However, it is not generally true for uniform gratings and will not be used in the rest of this chapter; and

- The multi-wave CWM of Kostuk neglects second order derivatives of field amplitude but still produces results as good as the RCWM of this chapter. This is only true for this case where all gratings are optically thick (see appendix A). In a general analysis of many gratings in reflection and transmission, recorded coherently or incoherently, the RCWM model is the more generally applicable.

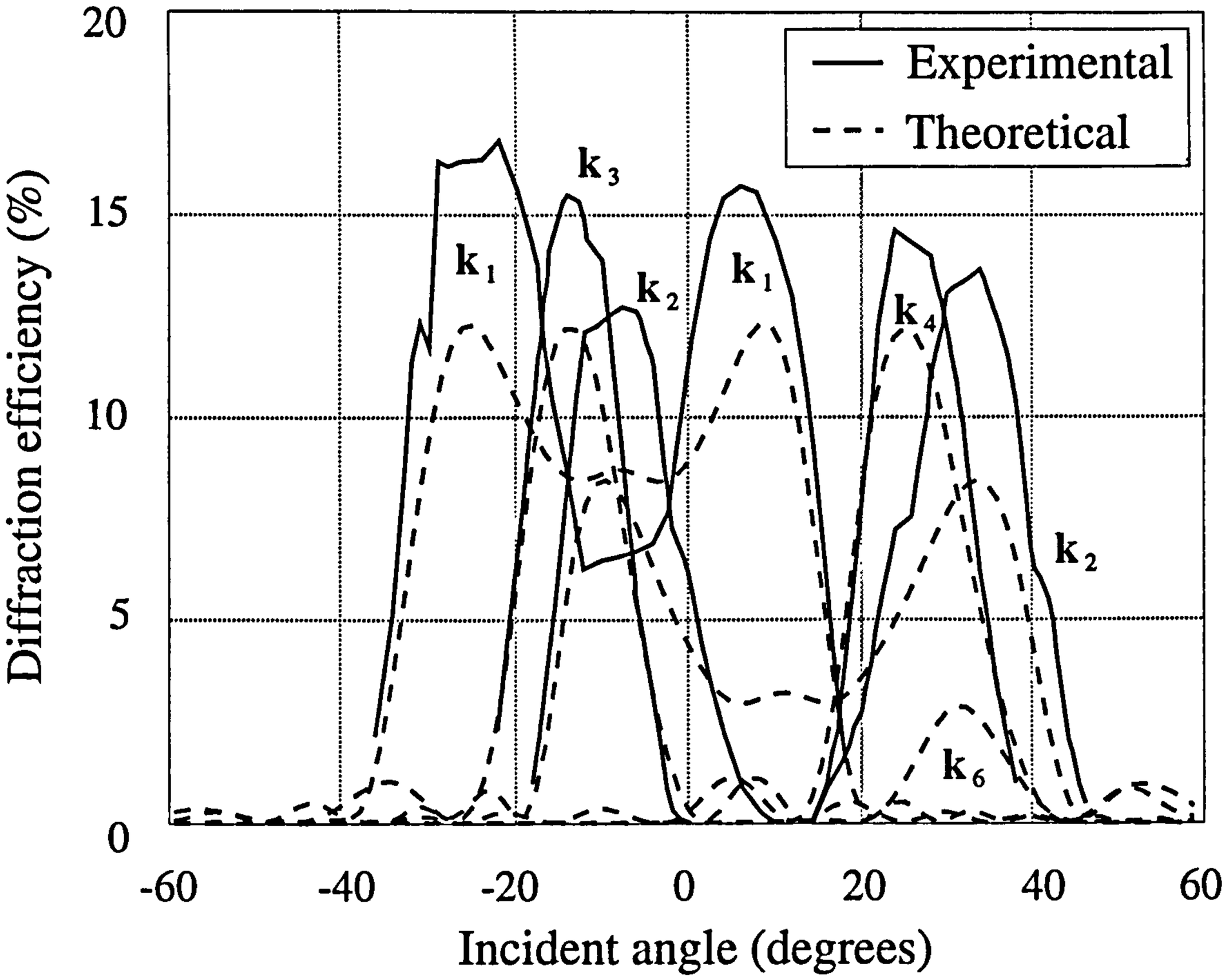


Figure 5.10: Theoretical RCWM comparison of results with the experimental data of Kostuk et al. [KOS86], for the *simultaneous* recording of a reflection fan-out to two hologram. The **k** vectors and coupling labelled above correspond to those shown in figure 5.8b.

The paper of Kostuk [KOS86] proceeds to analyse the sensitivity of reflection fan-out gratings to emulsion thickness changes. As with all reflection HOEs the expected replay fidelity is greatly affected by processing changes to the emulsion and these should be minimised. The RCWM model agrees with this analysis and it will be discussed further in chapter 6.

This section has established the credentials of the new RCWM model by demonstrating its agreement with published results. It will now be used to theoretically analyse reflection fan-out gratings in more detail.

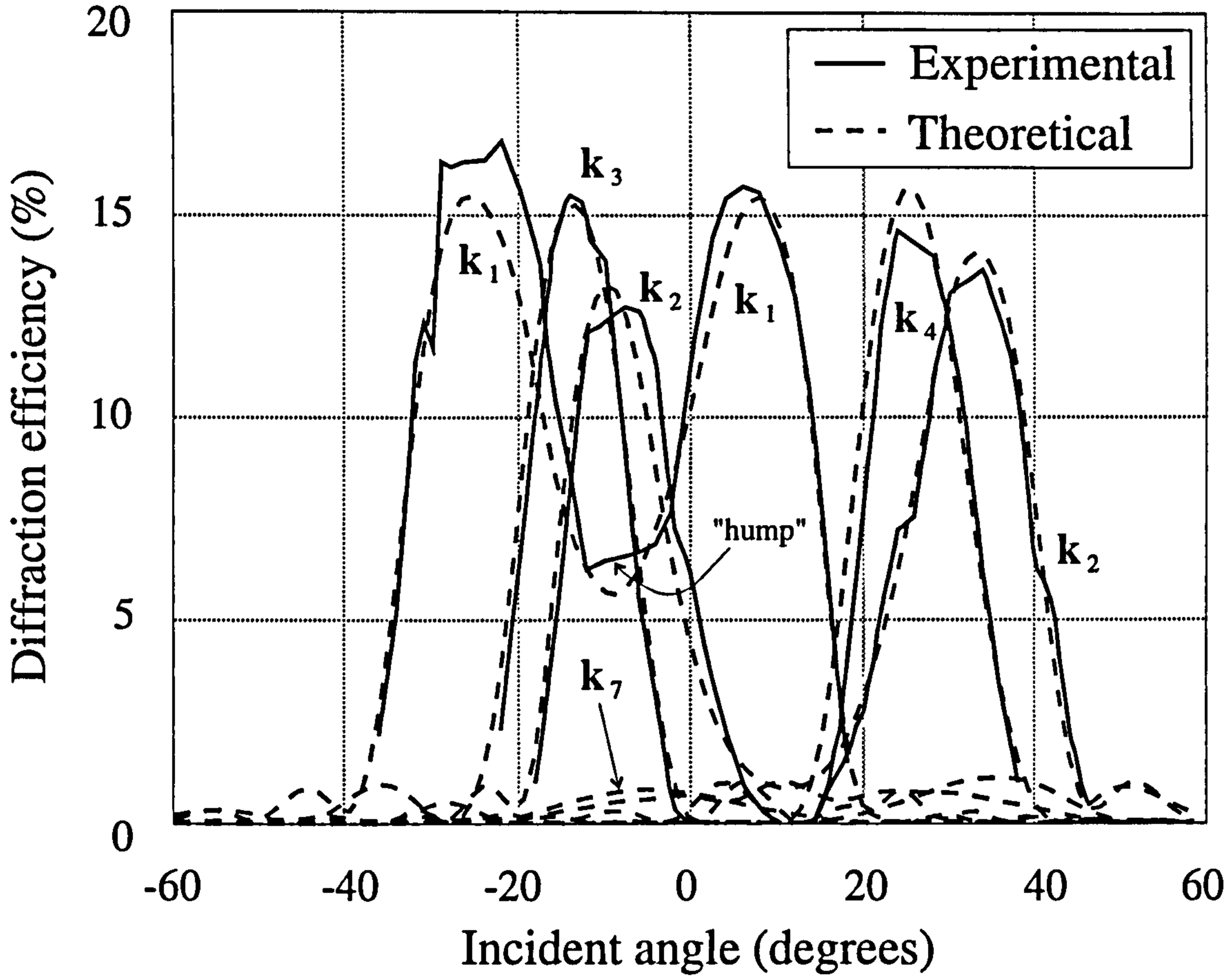


Figure 5.11: Theoretical RCWM comparison of results with the experimental data of Kostuk et al. [KOS86], for the *simultaneous* recording of a reflection fan-out to two hologram. The k vectors and coupling labelled above correspond to those shown in figure 5.8a, (those diffraction curves not labelled correspond to reflected secondary wave power).

5.7 Reflection fan-out analysis

The aim of this section is to give a more detailed analysis of reflection fan-out gratings, in particular pseudo-conformal gratings, with a single common replay beam angle. It is shown that high efficiency and a low uniformity error can be achieved in reflection using a simple optical recording process where grating phases do not need to be controlled.

Fan-out holograms with a separate reference beam for each object wavefront will usually have gratings that are Bragg mismatched/decoupled from one another and so very little interaction will occur between the gratings. In contrast, those with a common reference can have a degree of Bragg overlap between gratings that affects

the way in which they replay. This analysis concentrates on this the second class of multiple grating hologram.

The common recording reference beam angle of a reflection volume fan-out hologram, θ_R, (see figure 5.1) can be either on- or off-axis and still give a high diffraction efficiency, because the recording beams are counter-propagating, (transmission gratings have to be off-axis). An in-depth analysis is presented here for on-axis ($\theta_R = 0°$) gratings with relatively small inter-object-beam angles, $\delta \leq 8°$. The results and characteristics are broadly the same for off-axis gratings. On-axis gratings are of particular interest in optical interconnection systems because they match the optical axis of other optical components.

Reflection gratings with an on-axis reference beam and very small object beam angles have fringes very similar to a pure reflection grating (PRG) and as such can be called pseudo-conformal.

The RCWM model can deal with multiple gratings recorded incoherently or coherently simply by altering the coupling parameter χ_a, (equation 5.12), to include the secondary grating, a, recorded between two object beams. Different beam ratios (strengths between primary and secondary gratings) can be similarly handled.

5.7.1 Fan-out to two

In this analysis the following physical and recording parameters are chosen for the hologram: the gelatin thickness, $d = 15\mu$m, reference beam recording angle, $\theta_R = 0°$, recording and replay wavelengths are both 514.5nm and the average composite object beam angle is 0°, (typical experimental parameters for DCG HOEs derived from Kodak 649F emulsions). Bulk absorption, α, is assumed to be zero, refractive indices each side of the element are 1.0, bulk index of the gelatin is 1.5 (no changes to the emulsion occur during processing) and the gratings are replayed at the Bragg angle. The whole analysis is similar to that of Redmond [RED89a] for transmission holograms and a direct comparison can be made.

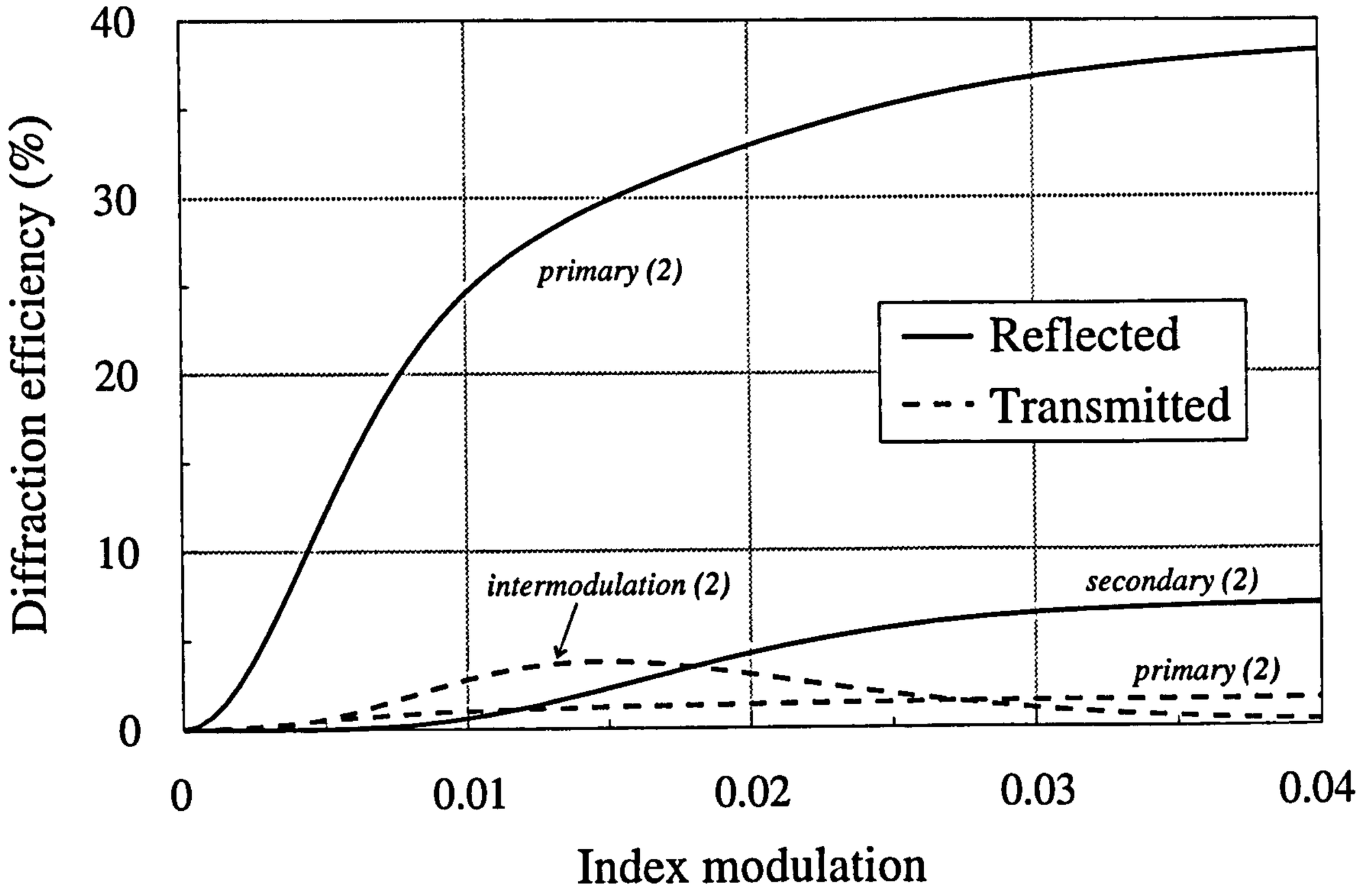


Figure 5.12: Diffraction efficiency vs. refractive index modulation for a fan-out to two pseudo-conformal reflection hologram. Incoherent recording with $\delta = 8°$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

Sequential (*or* Incoherent)

For a sequential fan-out to two hologram recorded in a linear medium with equal strength gratings $\chi_1 = \chi_2$ and $\chi_3 = 0$. The variation of diffraction efficiency with index modulation for three different values of δ, (equal to 8°, 2° and 0.2°) are shown in figures 5.12 to 5.14. These figures present both reflected and transmitted power for primary, I- and S-waves, (the number of space-harmonics after degeneracy has been eliminated is 7).

It is clear from these figures that the efficiency trend is the same, no matter what the fan-out angle: power is diffracted into the two primary reflected orders with equal efficiency, approaching a total asymptotic limit (for the pair) of $\sim 78\%$, ($\Delta n_p \sim 0.05$). As index modulation increases the angular acceptance, $\Delta\theta_B$, of the primary gratings increases leading to a higher degree of Bragg overlap between the two gratings, ($\Delta\theta_B$ is defined as the deviation from the Bragg angle at which η drops to zero for a single grating). A consequence of this is that S-wave power increases in a

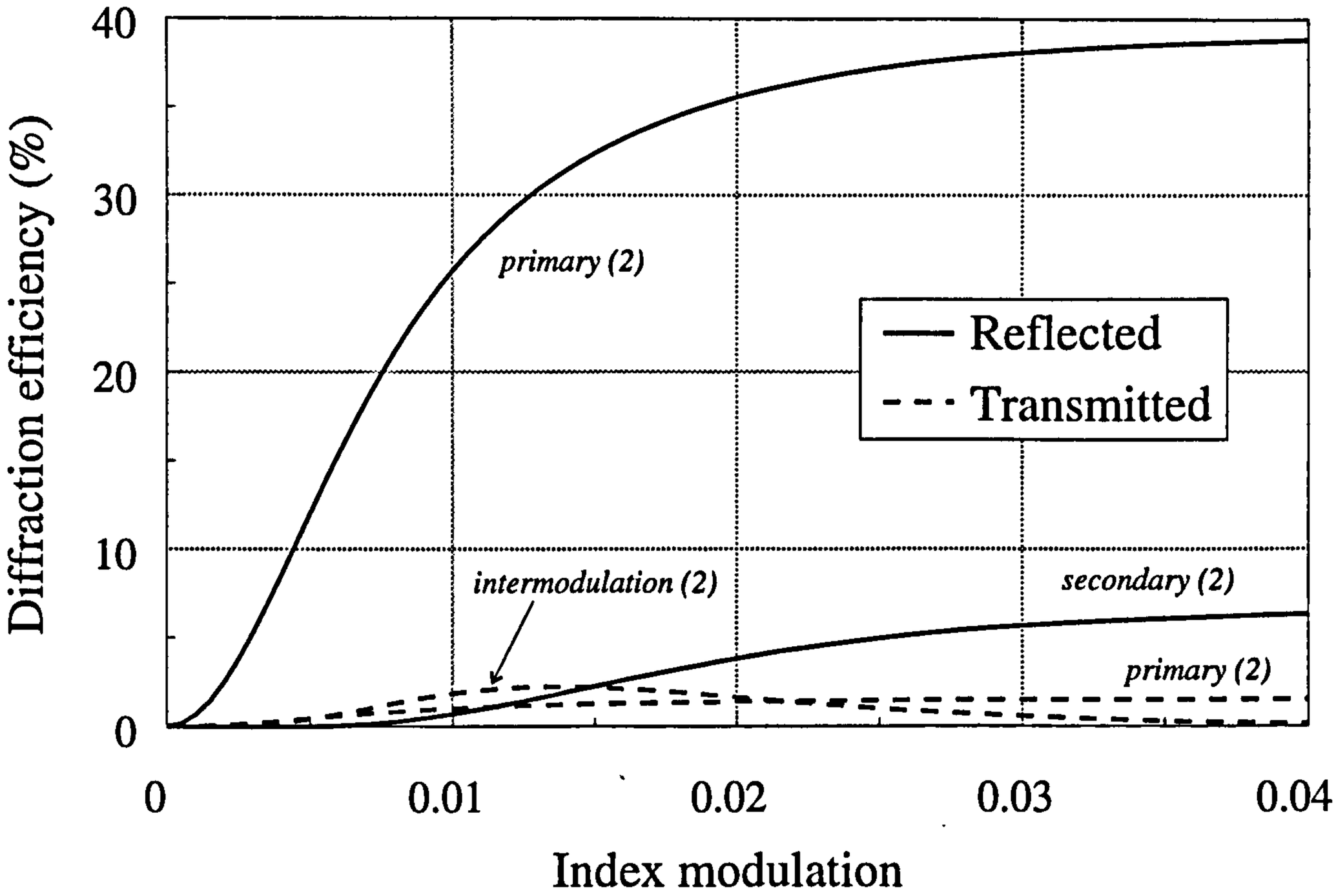


Figure 5.13: Diffraction efficiency vs. refractive index modulation for a fan-out to two pseudo-conformal reflection hologram. Incoherent recording with $\delta = 2°$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

similar way to primary power, however, since this coupling is via a double diffraction from the primary waves (in the absence of secondary gratings) the maximum power is limited to $\sim 18\%$ of the incident wavefront. The same argument can be applied to I-wave power: the zero-order (replay) wavefront is double diffracted from both primary gratings, however the index modulation where I-wave power peaks (~ 0.012) is lower than that of S-waves so that the coupling is weaker and the peak efficiency is much smaller ($< 3\%$).

Inclusion of the full 19 waves illustrated in figure 5.3 produces no significant change to these results. It is clear from the graphs that small δ is marginally better, in terms of primary efficiency and reduced S-wave power, than $\delta = 8°$, but only by $\sim 1\%$. This is in stark contrast to the transmission grating results of Redmond [RED89a] that show much greater variation in η and ΔU with a change in δ.

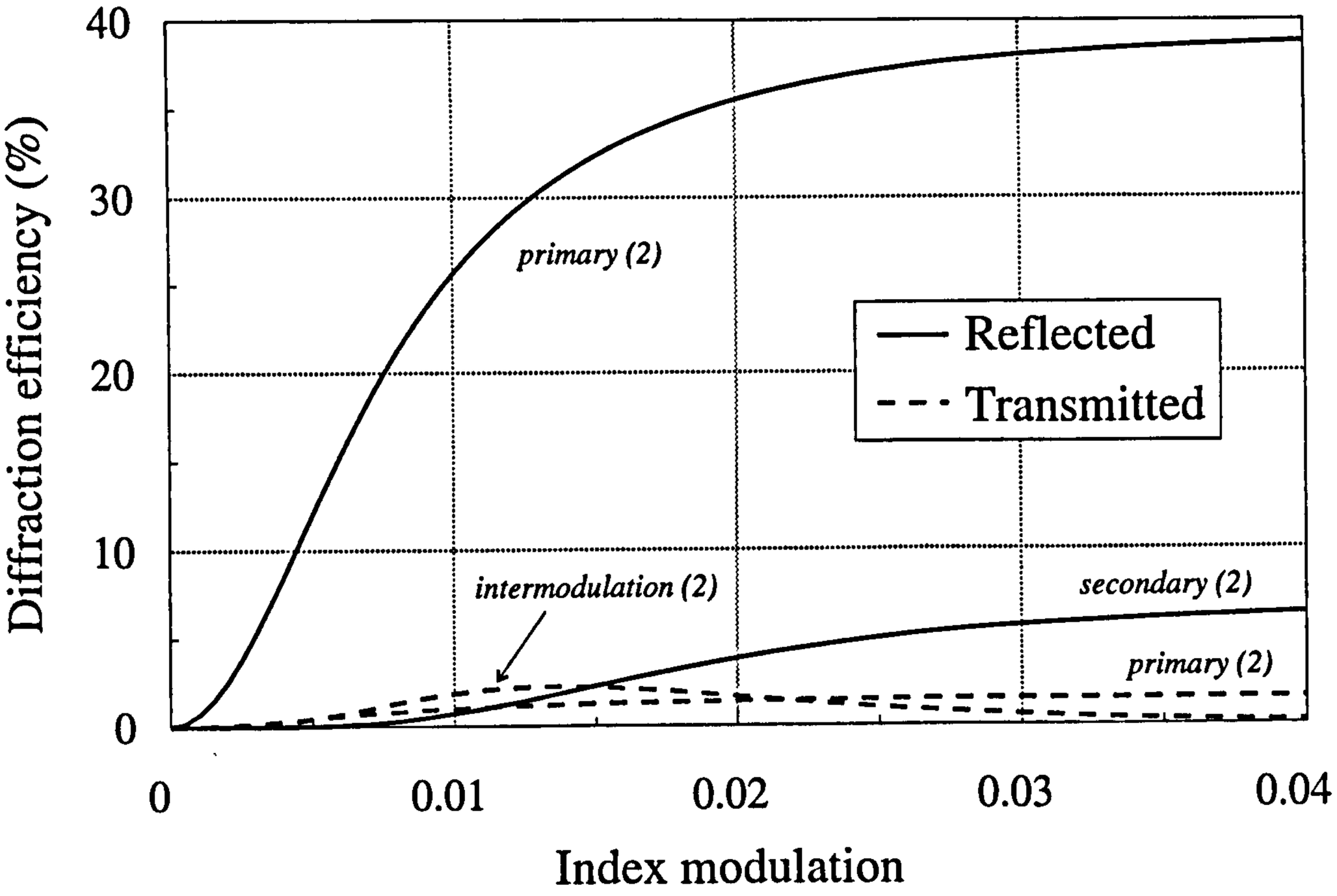


Figure 5.14: Diffraction efficiency vs. refractive index modulation for a fan-out to two pseudo-conformal reflection hologram. Incoherent recording with $\delta = 0.2°$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

Simultaneous (*or* Coherent)

A *real* secondary grating is now considered, recorded by two simultaneous object waves, the strength of which can be controlled by varying χ_3. The beam ratio, B (see section 5.3.2), is the experimental recording variable that is used to control the grating coupling parameter. The ratio of primary to secondary grating strengths, $\Delta n_p/\Delta n_s$, is quoted in this analysis, (rather than B). Varying the beam ratio can lead to a saturation of index modulation, however, consideration of this is left until section 5.7.3.

The variation of diffraction efficiency with index modulation is shown in figures 5.15 and 5.16 for primary and secondary gratings of equal strength, with $\delta = 0.02°$ and $0.2°$ respectively, (the number of space-harmonics is still 7). It is clear that the introduction of a secondary grating, equal in strength to the two primary gratings, has a significant effect on both primary wave efficiency and uniformity error. I-wave power is significant ($\sim 28\%$ of the total) at low values of index modulation

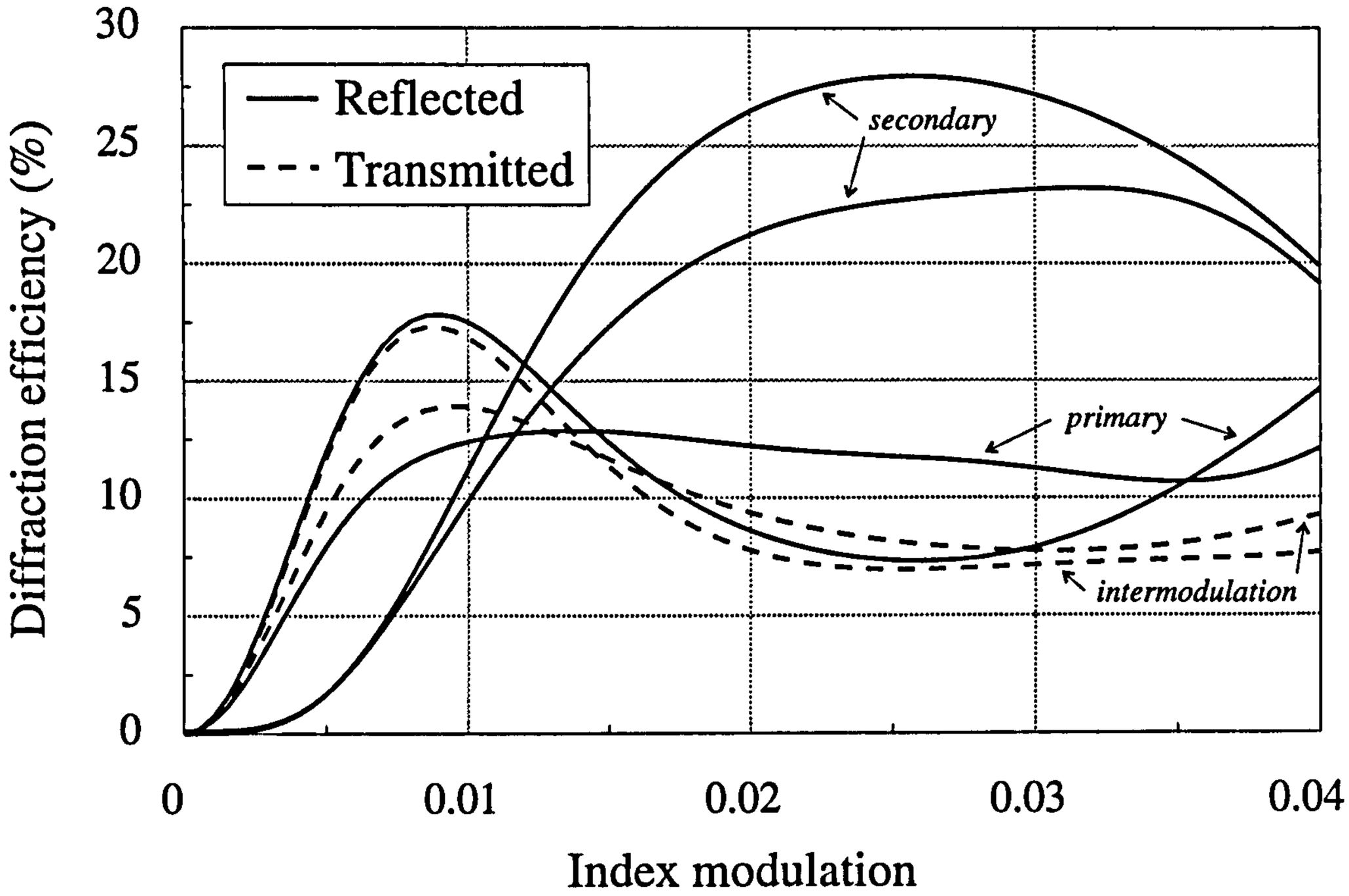


Figure 5.15: Diffraction efficiency vs. refractive index modulation for a fan-out to two pseudo-conformal reflection hologram. Coherent recording with $\delta = 0.02°$ and $\Delta n_p/\Delta n_s = 1$.

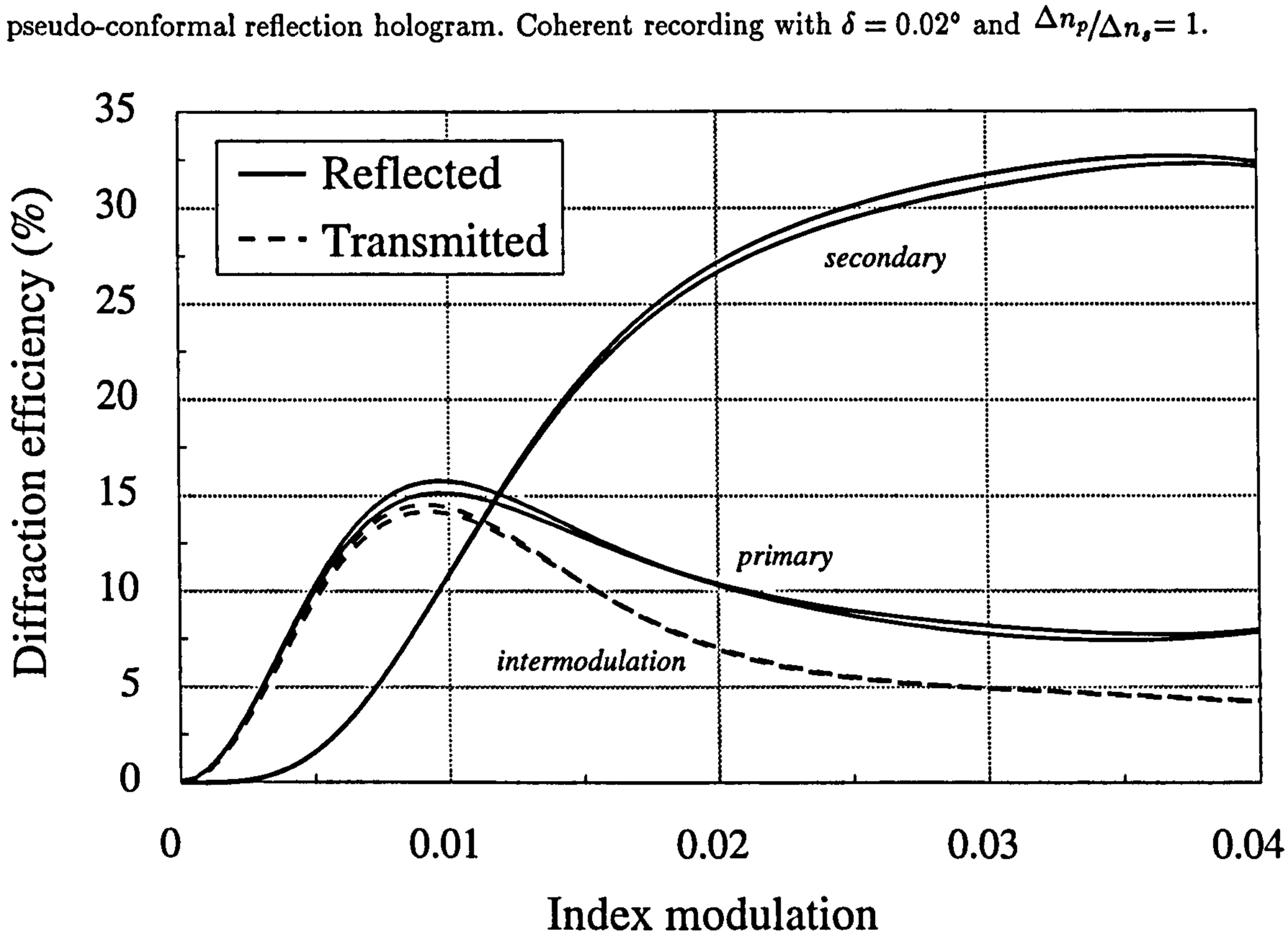


Figure 5.16: Diffraction efficiency vs. refractive index modulation for a fan-out to two pseudo-conformal reflection hologram. Coherent recording with $\delta = 0.2°$ and $\Delta n_p/\Delta n_s = 1$.

(~ 0.008) which is consistent with a relatively strong transmission grating coupling power out of the two primary diffracted directions. At higher index modulations the S-waves obviously become more optimally Bragg matched to the zero-order. This is due not only to a strong double diffraction from the primary waves via primary gratings, but also to a relatively strong coupling of power away from the primary diffraction orders via secondary gratings and from I-waves into the S-waves via primary gratings, (see figure 5.8b for coupling to k_5 & k_6). There is a clear increase

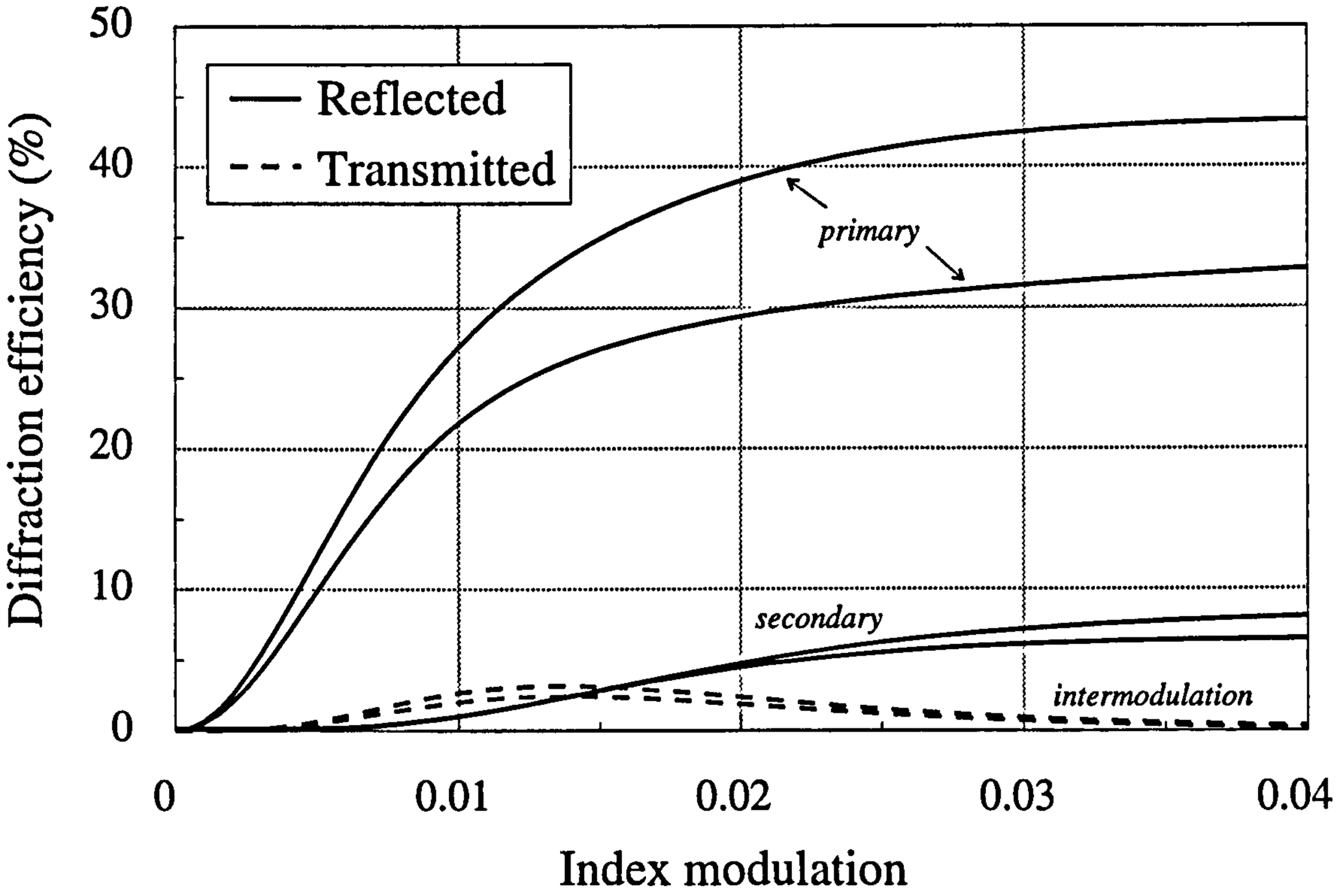


Figure 5.17: Diffraction efficiency vs. refractive index modulation for a fan-out to two pseudo-conformal reflection hologram. Coherent recording with $\delta = 0.02°$ and $\Delta n_p/\Delta n_s = 10$.

in the uniformity error of all waves (from figure 5.15) in the very small δ regime ($\delta = 0.02°$). The reason for this appears to be that there is a strong Bragg overlap between "thick" primary gratings at very small object beam separations. Consequently, secondary gratings for $B \sim 1$ are also thick and coupling takes place via the -1 grating order (and not ± 1 for "thinner" gratings). This leads to an asymmetric coupling of power between all orders and the resultant non-uniformity in diffraction efficiencies.

Figures 5.17 and 5.18 show the same grating geometry for a higher beam ratio

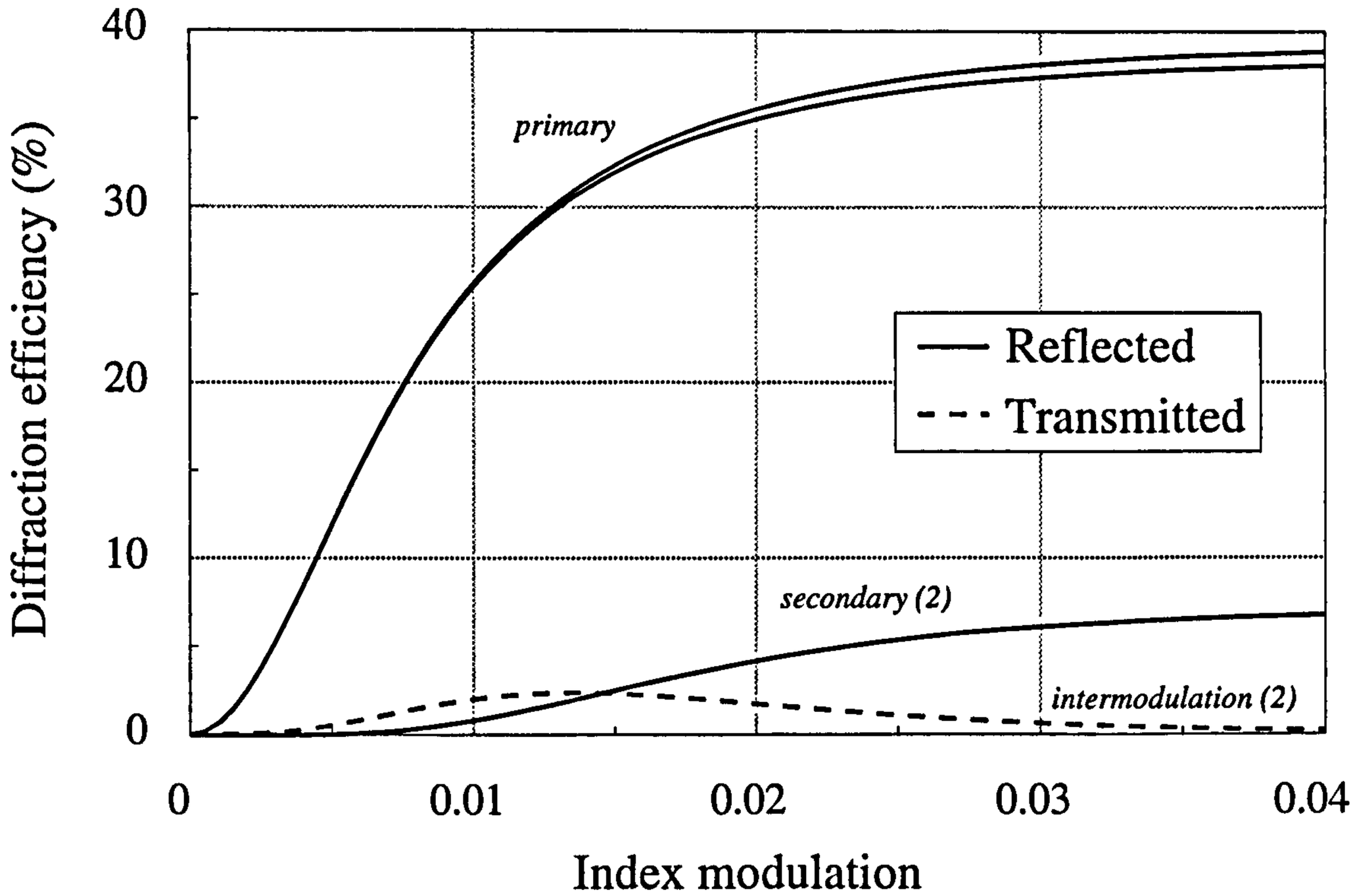


Figure 5.18: Diffraction efficiency vs. refractive index modulation for a fan-out to two pseudo-conformal reflection hologram. Coherent recording with $\delta = 0.2°$ and $\Delta n_p/\Delta n_s = 10$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

and reinforce the trends shown in previous figures. The higher beam ratio increases the efficiency and uniformity error of the primary waves towards the level of the incoherent recording case (i.e. $\eta \sim 78\%$ and $\Delta U = 0\%$). Evidence of asymmetric coupling is still present for the small δ ($\sim 0.02°$) case.

The conclusions to be drawn from the fan-out to two case are:

- η and ΔU for primary diffracted orders are at an optimum for incoherent recording or when coupling via secondary gratings is at a minimum,
- η and ΔU exceed those for transmission gratings [RED89a, pages 97 to 102]; and
- For coherent recording in the very small δ regime, asymmetric coupling is evident that leads to an increase in ΔU.

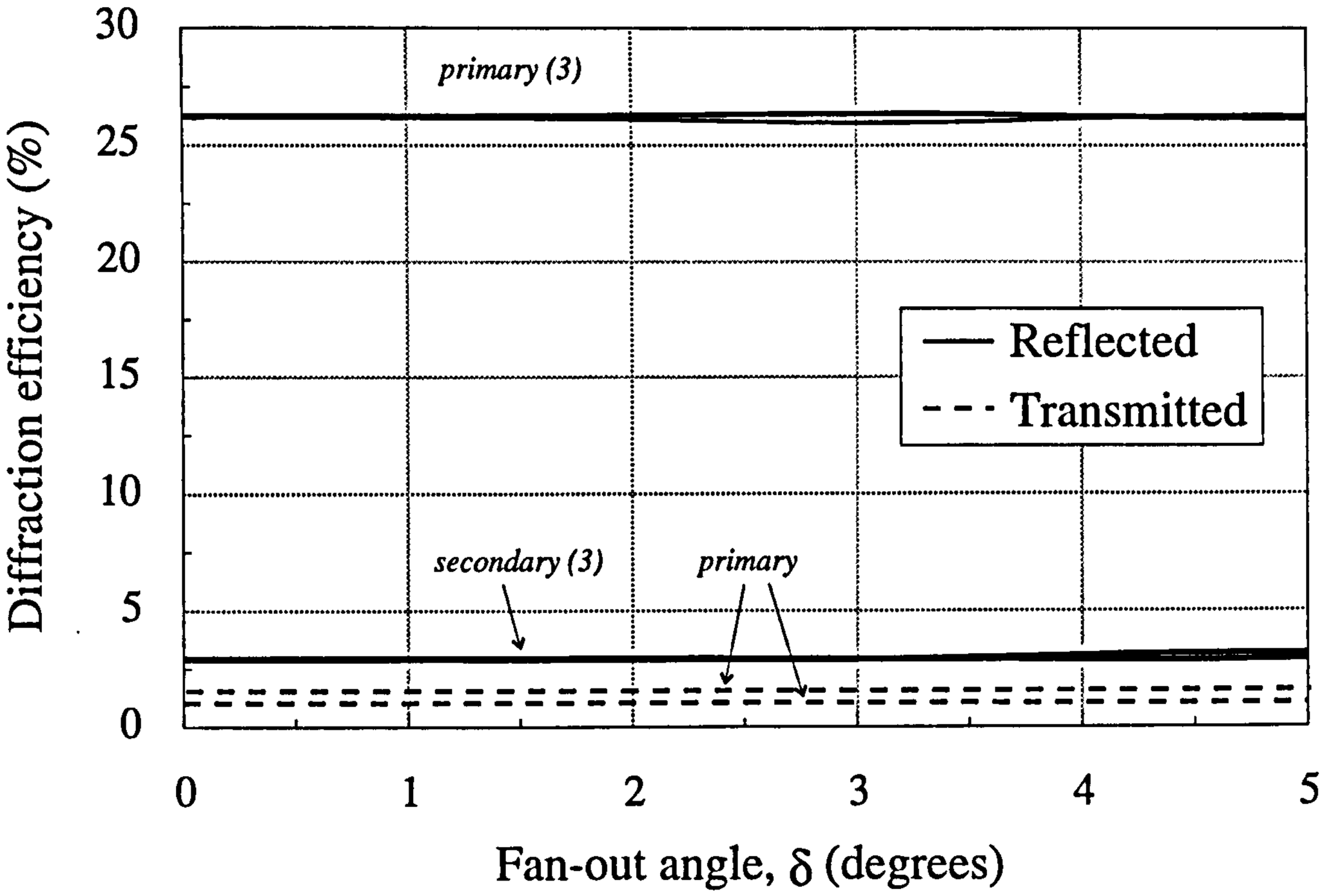


Figure 5.19: Diffraction efficiency vs. fan-out angle for a fan-out to three pseudo-conformal reflection hologram; incoherent recording $\Delta n_p/\Delta n_s = \infty$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

5.7.2 Fan-out to higher numbers

In the analysis of the last section the best efficiency and uniformity errors were achieved for index modulations ~ 0.05. Such modulations are unrealistically high for 15μm-thick gelatin, but lower modulations in 40μm-thick gelatin can produce the same results. However, for consistency with section 5.7.1 and to assist in a direct comparison with similar work on transmission gratings [RED89a, chapter 4] this theoretical analysis will continue assuming such high modulations (without saturation) in 15μm emulsions - with the caveat that thicker emulsions are required to realise such holograms in practice.

The following sections analyse the effect on η and ΔU by varying δ for higher levels of fan-out, (assuming a fixed modulation of primary and secondary gratings). All physical and recording parameters of the hologram are assumed to be the same as section 5.7.1. Primary grating modulation, Δn_p, is held at 0.05 in all cases, which, as the level of fan-out increases, represents a highly over-modulated grating;

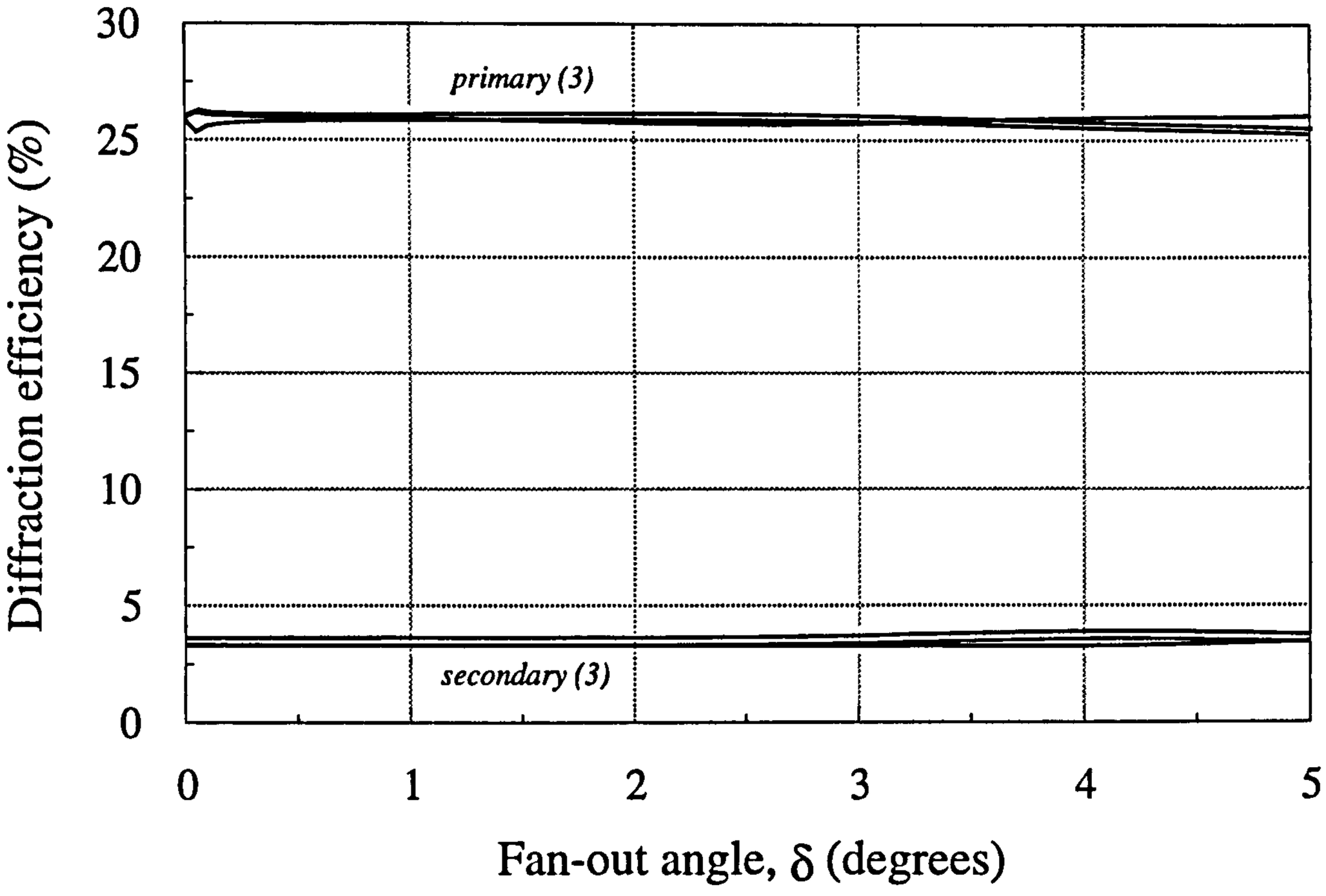


Figure 5.20: Diffraction efficiency vs. fan-out angle for a fan-out to three pseudo-conformal reflection hologram; coherent recording $\Delta n_p/\Delta n_s = 10$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

section 5.7.3 discusses the optimum Δn_p for N gratings.

Fan-out to three

Figures 5.19 to 5.22 show the variation of η with δ for four separate grating strength ratios, $\Delta n_p/\Delta n_s$. Where several power spectra are close to one another the figures in brackets state how many waves are present in the same region of the graph. Waves with efficiencies less than $\sim 1\%$ of the total power are not shown, (the total number of non-degenerate waves is 19).

Clearly the trend is similar to the two grating situation. Efficiency and uniformity are best for $\Delta n_p/\Delta n_s = \infty$ and 10. At lower beam ratios secondary wave power increases significantly, so that when $\Delta n_p/\Delta n_s = 1$ the power in the primary waves, although having good uniformity over δ, is only $\sim 6\%$ of the total. For coherent recording, the same uniformity error at very small δ is present.

There is a potential to obtain a higher η and low ΔU for the situation

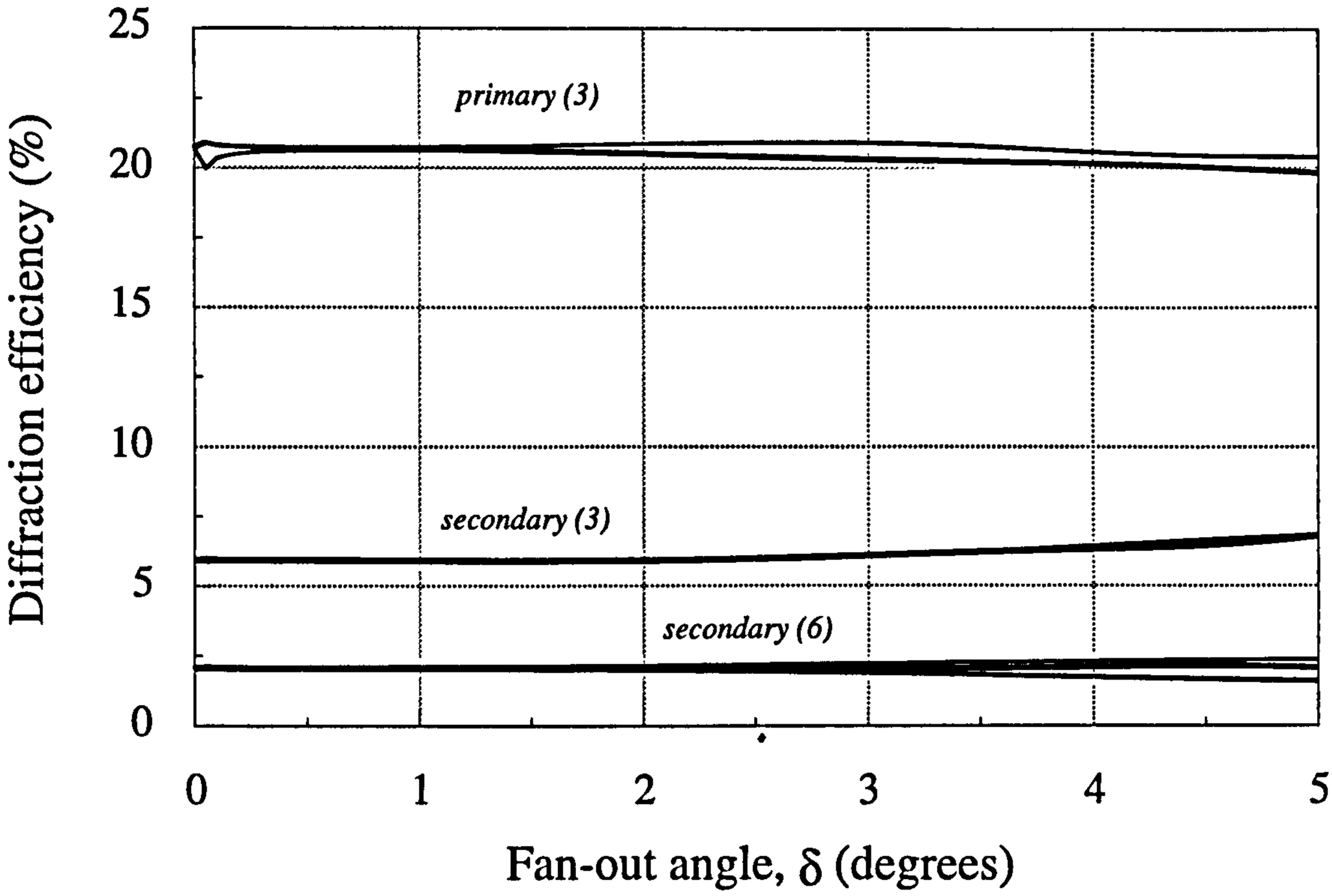


Figure 5.21: Diffraction efficiency vs. fan-out angle for a fan-out to three pseudo-conformal reflection hologram; coherent recording $\Delta n_p/\Delta n_s = 3$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

where $\Delta n_p/\Delta n_s = 1$ by designing the HOE such that the three desired output wave directions correspond to the three highest efficiency secondary waves of figure 5.22. This would still limit useful $\eta \sim 36\%$ which is not a very satisfactory solution.

Fan-out to four

Figures 5.23 to 5.26 replicate the fan-out to three analysis, but for four primary gratings. The results reflect the fan-out to three case, so only primary wave reflected power is shown.

It can be seen, from the three and four grating cases, that $\eta \sim 60\%$ and $\Delta U \sim 0\%$ are observed for a range of inter-object beam angles where $\Delta n_p/\Delta n_s \geq 3$. This is far better than the comparable performance of transmission gratings analysed by Redmond [RED89a, pages 105 to 111]. The discrepancy between these separate analyses is most evident for $\Delta n_p/\Delta n_s \sim 1$, when it is clear that the inclusion of S-waves is important as they severely restrict primary wave diffraction efficiencies.

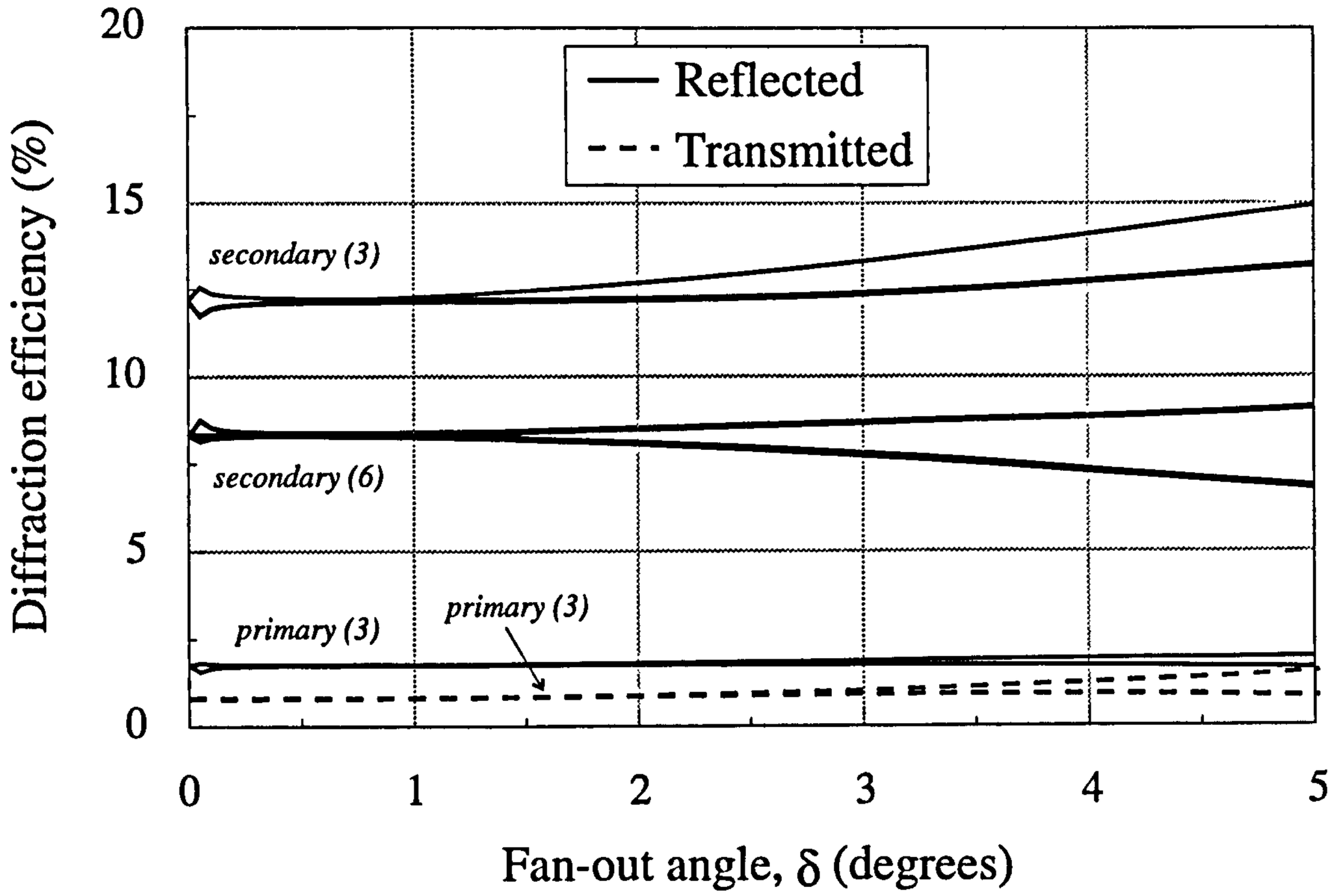


Figure 5.22: Diffraction efficiency vs. fan-out angle for a fan-out to three pseudo-conformal reflection hologram; coherent recording $\Delta n_p/\Delta n_s = 1$, (figures in brackets show the number of diffracted orders represented in the same region of the graph).

Fan-out to many

Figures 5.27a and 5.27b show that the maximum useful (primary wave) diffraction efficiency versus the degree of fan-out for incoherently recorded gratings in transmission and reflection with $\delta \approx 0°$. These are real efficiencies including Fresnel losses derived using the RCWM model presented here. The experimental transmission grating results of Redmond [RED89a, page 111] are shown in figure 5.27a; they have been modified to allow for Fresnel losses that were removed from the original data. The RCWM theoretical results correlate remarkably well with the experimental results, placing an upper limit on efficiency of $\sim 48\%$. It is clear from figure 5.27 that reflection fan-out achieves better results than transmission, exhibiting no drop-off in realisable efficiency ($\eta \sim 78\%$) with an increase in the level of fan-out, provided that enough index modulation is available.

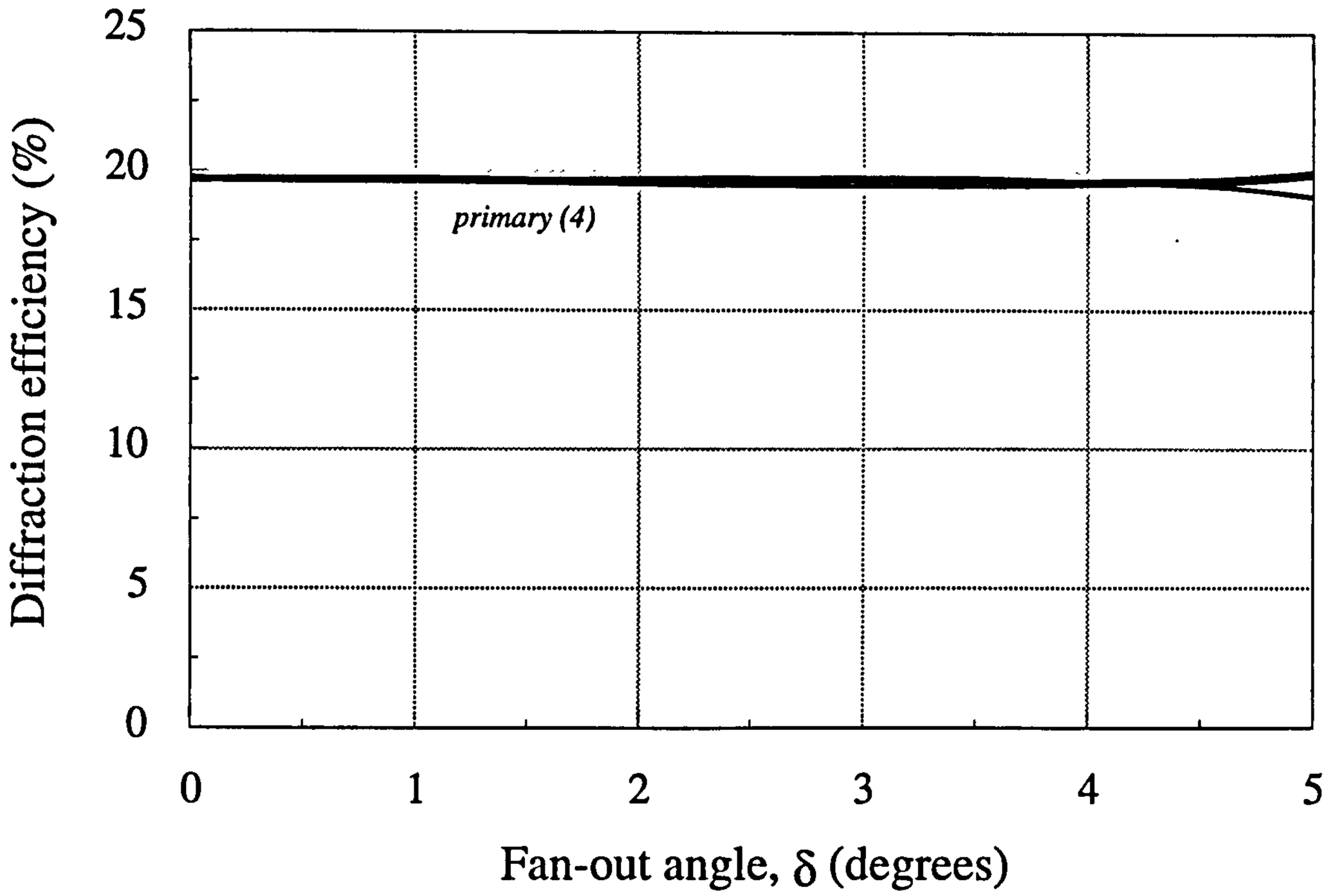


Figure 5.23: Diffraction efficiency vs. fan-out angle for a fan-out to four pseudo-conformal reflection hologram; incoherent recording $\Delta n_p/\Delta n_s = \infty$.

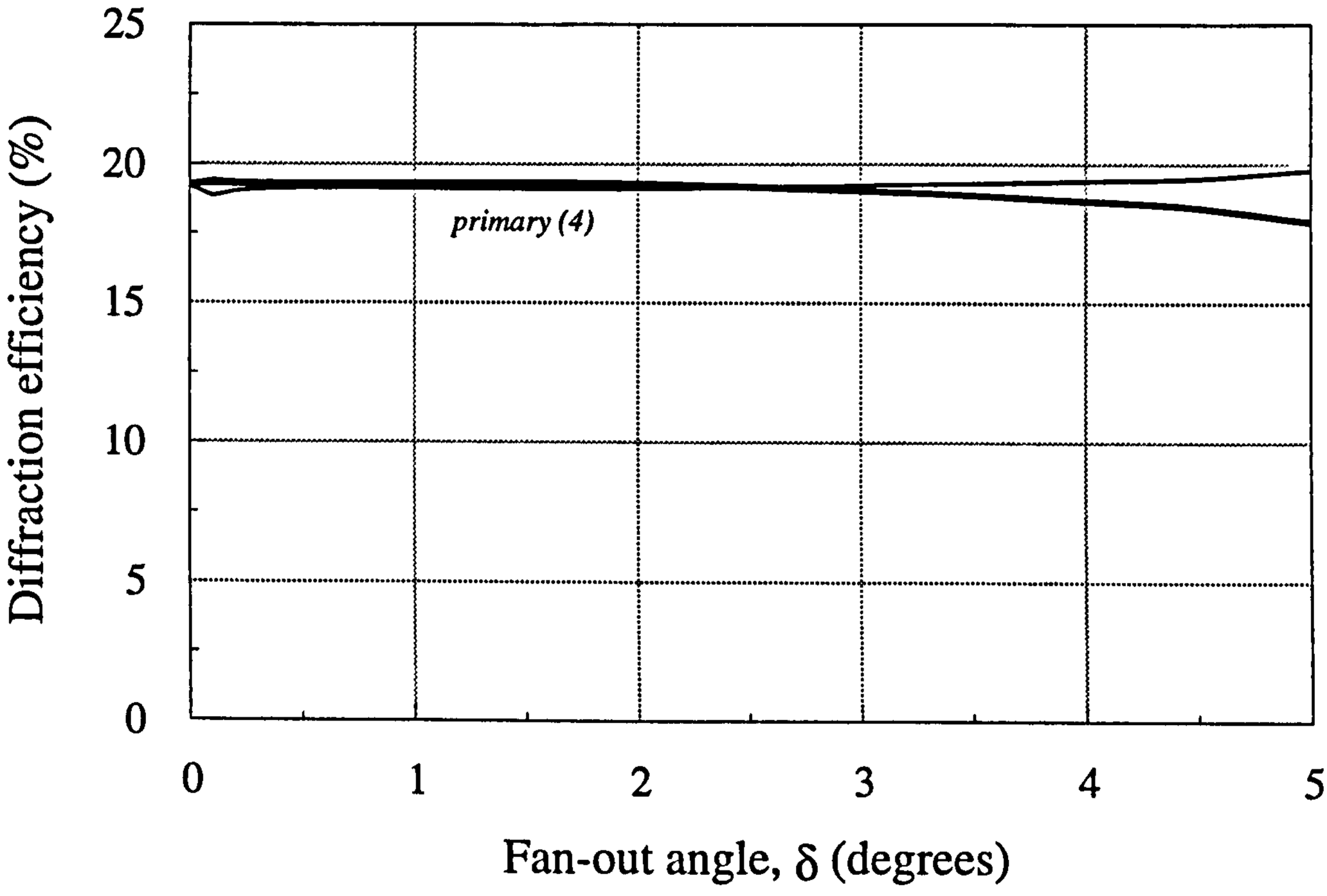


Figure 5.24: Diffraction efficiency vs. fan-out angle for a fan-out to four pseudo-conformal reflection hologram; coherent recording $\Delta n_p/\Delta n_s = 10$.

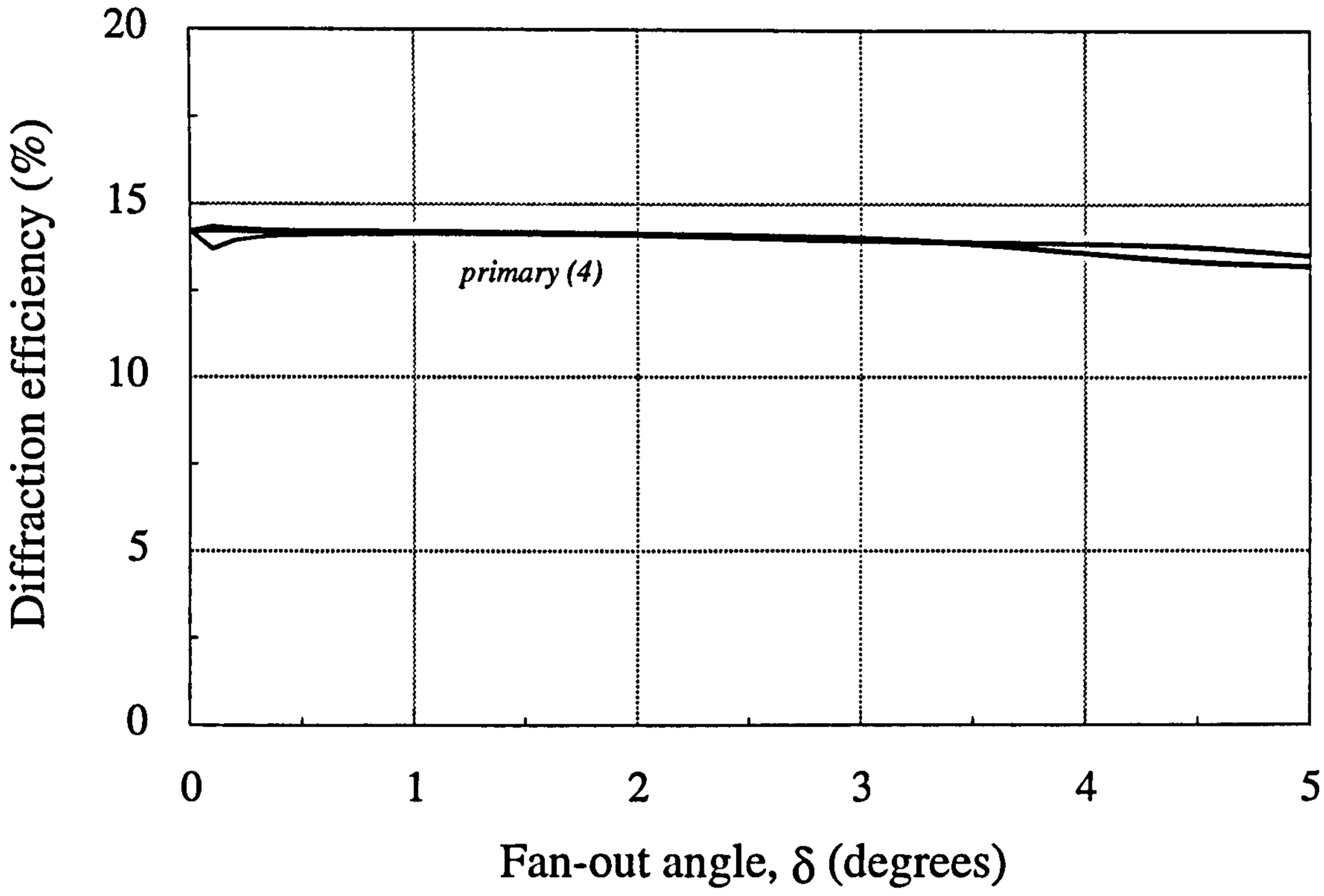


Figure 5.25: Diffraction efficiency vs. fan-out angle for a fan-out to four pseudo-conformal reflection hologram; coherent recording $\Delta n_p/\Delta n_s = 3$.

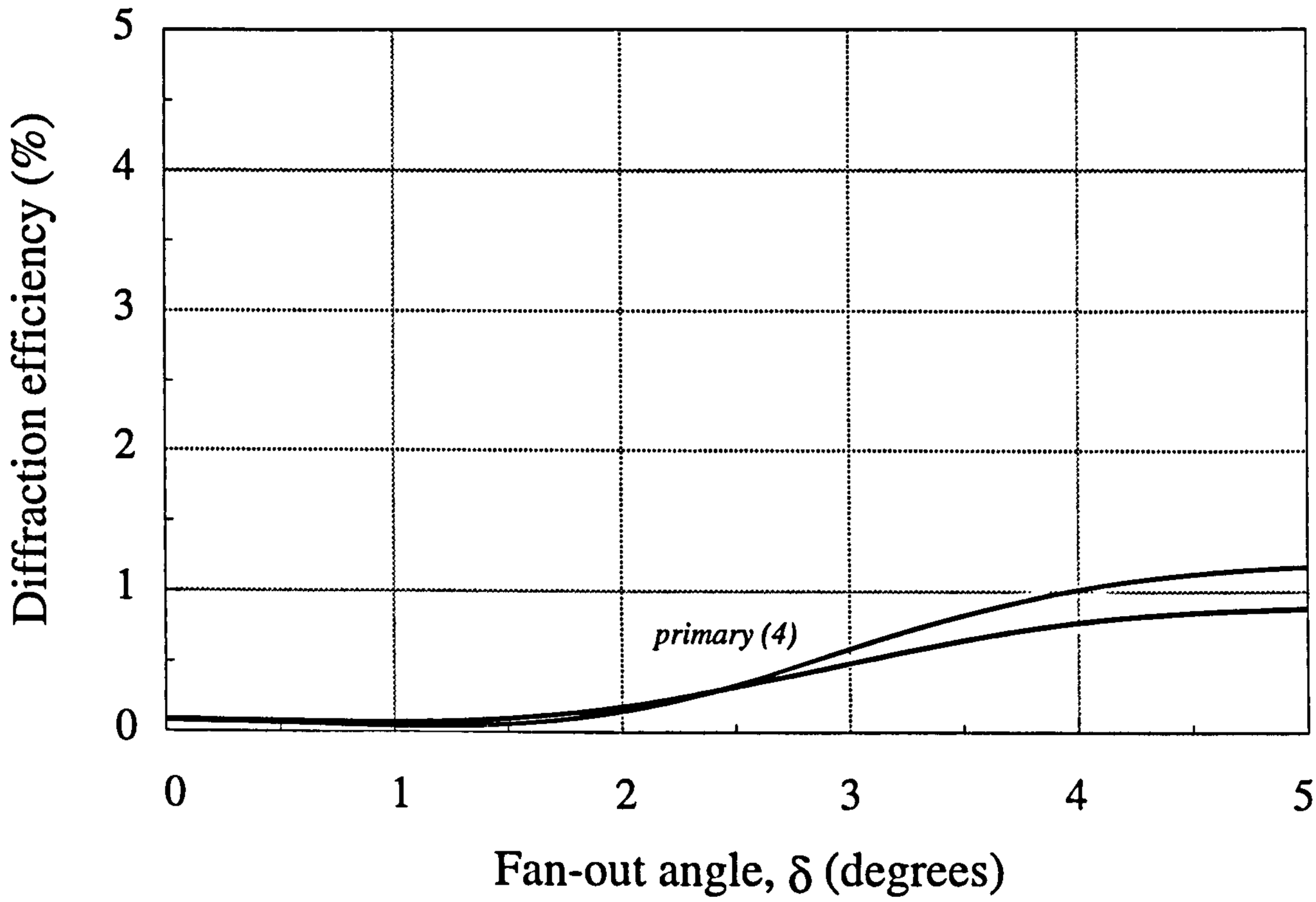


Figure 5.26: Diffraction efficiency vs. fan-out angle for a fan-out to four pseudo-conformal reflection hologram; coherent recording $\Delta n_p/\Delta n_s = 1$.

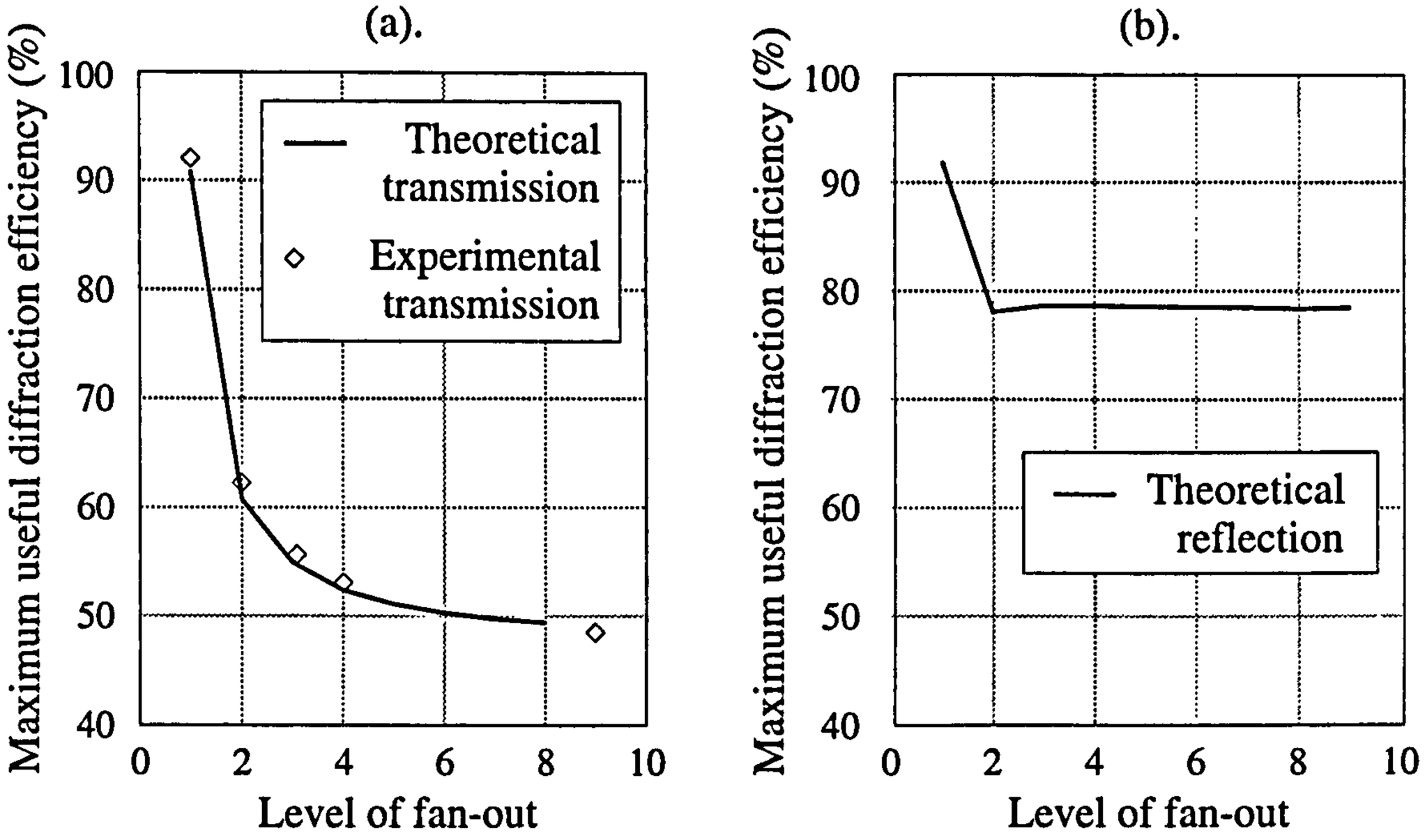


Figure 5.27: Theoretical results for maximum diffraction efficiency versus degree of fan-out for $\delta \approx 0$: (a). Transmission fan-out; and (b). Reflection fan-out gratings. Transmission results show the experimental data of Redmond [RED89a, page 111] modified for Fresnel losses.

5.7.3 Saturation effects

DCG, as with all holographic media has a limited dynamic range. Chapter 3 includes a discussion of the effects of saturation of the refractive index modulation for a single grating, where harmonic grating frequencies are produced. When two or more gratings are multiplexed in the same volume, saturation causes sum and difference frequency gratings as well as the harmonics.

In addition, the total modulation required to achieve peak efficiency for N gratings is $\sqrt{N}$ times that for a single grating [CAS75, page 726], making saturation more likely.

In an ideal linear medium with no saturation, an incoherent recording of multiple object waves does not produce intermodulation gratings. However, LaMacchia & Vincelette [LAM68] discuss a major drawback to this recording technique. The essence of this is that for the same total exposure the diffraction efficiency of an incoherently recorded N grating fan-out HOE is N times less than the coherently recorded counterpart, (they assume a high beam ratio in the coherent case so

that cross-grating interaction is negligible). Since the recorded grating structure is defined by the irradiance distribution in the hologram plane, the refractive index modulation will depend upon $E_{seq}(\mathbf{r})E^*_{seq}(\mathbf{r})$ and $E_{sim}(\mathbf{r})E^*_{sim}(\mathbf{r})$ for the incoherent and coherent recording situations respectively, (see section 5.3.1 equations 5.1 and 5.2). The irradiance distribution for sequential recording has a DC background component N times higher than the simultaneous case. Consequently, for peak diffraction efficiency in both cases the sequential recording uses up the available index modulation more quickly, leading to saturation.

To avoid saturation, a thicker gelatin is required to efficiently record high levels of fan-out. This is particularly true for pseudo-conformal reflection gratings where the index modulation required to achieve peak efficiency is higher than for transmission gratings of the same thickness. Generally, for fan-out to N, it is the product Δnd, (where d is the emulsion thickness) that must be $\sqrt{N}$ times greater than the single grating value for peak efficiency.

Redmond [RED89a, pages 112 to 123] performed a detailed analysis of the effects of saturation on diffracted directions, efficiency and uniformity of modulation. The RCWM of this chapter can be extended to include harmonic, sum and difference frequency gratings in order to model saturation. Alternatively, the following argument can be applied. For sequentially recorded gratings in a saturating medium, harmonic and sum gratings are usually far from Bragg incidence and only have a small effect. The difference frequency gratings are equivalent to the secondary gratings present in a simultaneously recorded hologram. Therefore a saturated incoherently recorded fan-out HOE can be approximately modelled by a coherently recorded element.

The overall effect of saturation is undesirable: coupling power into unwanted diffracted directions, increasing the uniformity error and decreasing useful diffraction efficiency. Consequently, for DCG fan-out interconnects, thick gelatin layers $\geq 40\mu$m are desirable.

5.7.4 Analysis conclusions and model improvements

In contrast to equivalent transmission elements, reflection fan-out HOEs can, in theory, produce high efficiency and a low uniformity error independently of the size of the angular separation of the object waves, δ, (provided that the medium is thick enough or the available index modulation is large enough). Redmond [RED89a] defined two useful efficiency regimes for transmission elements, in terms of the angular acceptance, $\Delta\theta_B$: $\delta/\Delta\theta_B > 1$ and $\delta/\Delta\theta_B << 1$. In the first regime gratings are closely coupled and in the second they can be treated as being decoupled, when replayed at the Bragg angle. For a transmission grating, (recorded in a fixed thickness of emulsion), $\Delta\theta_B$ decreases as Δn increases, so at peak primary grating efficiency $\Delta\theta_B$ is small [KOG69]. In contrast, $\Delta\theta_B$ for a reflection grating increases as Δn increases, so that at peak efficiency $\Delta\theta_B$ is always relatively large for a volume hologram. In addition, $\Delta\theta_B$ can be increased further by over modulating the grating at no expense to η, (provided the medium doesn't saturate). Hence, $\delta/\Delta\theta_B << 1$ over a large range of δ for a reflection fan-out element.

For a coherent recording, the cross-gratings in a reflection element are usually not at peak strength because they are transmissive. Therefore the affects of multiple diffraction: limiting η and increasing ΔU, are not as pronounced as they are in transmission elements except for relative index modulations ($\Delta n_p/\Delta n_s$) less than 3.

With index modulations ~ 0.028 (per grating) required to achieve peak efficiency, for the moderate level of fan-out to four, in a 15μm (emulsion thickness) pseudo-conformal hologram, it is clear that thicker gelatin emulsions are needed to produce practical HOEs of this type. However, thicker emulsions are encumbered with their own drawbacks due, principally, to the wet processing required when using DCG. Grating non-uniformities [NEW87, RED89a] inevitably affect the efficiency and replay fidelity of such elements, in a way that cannot be predicted by the RCWM used in this chapter. Consequently, it seems that only thick DuPont [SMO89] photopolymer or photorefractive crystals are appropriate materials in which to realise these reflection HOEs, because they do not require the wet processing that is responsible for grating non-uniformities in thick DCG emulsions.

Adding grating phases

For coherently recorded transmission fan-out elements, close to 100% efficiency can be achieved with a low uniformity error, $\Delta U \sim 0\%$, by controlling the relative phases of primary gratings (during recording) over a limited depth of emulsion [HER92, HER93]. This minimises the spatial beating between superimposed gratings [RED89a, HER93] and all but eliminates the presence of cross-gratings [HER92, page 5721].

An additional benefit of this phase control is that for coherent recording the index modulation required for peak efficiency is independent of the level of fan-out and for incoherent recording the $\sqrt{N}\Delta nd$ requirement is significantly reduced [HER93, page 201]. This reduction in the required dynamic range would be of additional benefit with reflection holograms where that required for peak efficiency is significantly larger than transmission equivalents.

The addition of grating phases to this model should be relatively straightforward, since it has already been done with the single grating RCWM by Zylberberg & Marom [ZYL83]. Actually calculating optimum phases for large levels of fan-out [RED89a, pages 120 to 121] and accurately fabricating reflection elements (which are harder to record than transmission elements, see section 2.9) is more problematic.

Conclusions on the advantages of reflection fan-out

A significant advantage of reflection fan-out over their transmission counterparts is that the intermodulation gratings are transmissive in nature and hence I-waves propagate out of the back of the hologram in the opposite direction to the primary diffracted orders. In addition, the index modulation that produces optimum efficiency for a reflection (primary) grating will not usually coincide with that for a transmission (secondary) grating. Hence, secondary gratings will not be as strong as primary and so in many instances the coupling from primary to secondary waves will also be weak, resulting in a higher signal diffraction efficiency and replay fidelity, (uniformity).

5.8 Summary

A novel extension to the RCWM of Moharam & Gaylord has been presented together with its use to mathematically analyse generalised fan-out elements. Theoretical results for multiple transmission and reflection phase gratings have been shown to match experiment and other theories. In addition, it has been shown that both S-waves and I-waves must be retained in a fan-out analysis when secondary grating modulation approaches that of the primary gratings, i.e. coherent recording.

The key result of this chapter is to show that *theoretically* a high useful diffraction efficiency, ($\sim$ 78%), and a low uniformity error, ($<$ 1%) can be achieved using pseudo-conformal reflection gratings without resorting to the grating phase control required to achieve similar results in transmission fan-out [HER92, HER93]. This reinforces the conclusions of Kowarschik [KOW78b] for the two grating case.

Chapter 6

Spatially Invariant Interconnects

The definition of spatial invariance, applied to optical interconnects, is described in section 1.3.2. This chapter concentrates on three interconnects that can be considered to be space-invariant, although space-variant alternatives are possible. These space-invariant interconnects (SIIs) are:

1. **Planar interconnect** - a light guiding glass plate is used as a 'local bus' board-level communication system for optoelectronic integrated circuits (ICs).

2. **Nearest neighbour interconnect** - a space-invariant optical system is used to connect each pixel in an interconnected logic array (conforming to the model of figure 1.2) to its immediate neighbours; and

3. **Crossbar switch interconnect** - an optical vector-matrix multiplier implementation is discussed, the optics of which, when considered separate to the switching elements (in this case a spatial light modulator), perform the same function on each of the inputs to produce the output. This is used primarily as an example of how multi-faceted volume holographic lenslet arrays (with a common Bragg angle) can be used in space-invariant systems.

The optical interconnection systems are briefly described for all three examples, but details and results are only presented for the volume holographic components; being the elements of interest in this thesis.

6.1 Planar interconnects

A variation on free-space interconnects that is of particular interest for connecting optoelectronic chips together (a 'local-bus' at the circuit board level) is the planar interconnect (sometimes called a substrate mode HOE). Such interconnection schemes utilise a light guiding glass plate to send signals laterally between optoelectronic transmitters and receivers on an electronic processing board. Planar interconnects have also been proposed for use as optical backplanes or in board-to-board communication systems [HAU91]. An idealised light-guiding plate system is shown in figure 6.1. This interconnect overcomes many of the bottlenecks inherent in board-level electrical connections, (discussed in chapter 1). However, this is not the only optical solution to the problem, other interconnects have been proposed [KOS85, CRA92, PLA95].

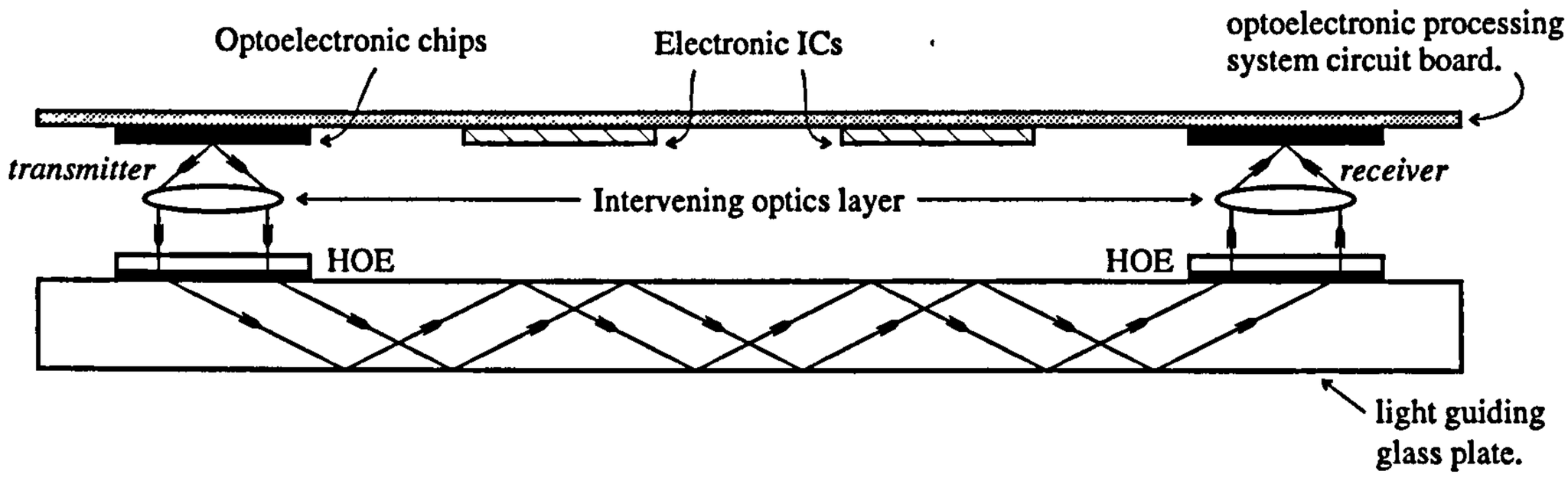


Figure 6.1: Conceptual view of board level chip-to-chip interconnects using a light guiding plate and illustrating the use of volume hologram coupling elements.

A key component in planar optical bus interconnection systems is an efficient holographic coupling element, used to transfer the light signals into the glass plate at an angle beyond the critical angle for the glass. The optoelectronic board containing processing chips, laser sources, modulators and detectors lies immediately above the light guiding plate, as shown. Communication between chips can occur via arrays of optical channels (parallel) or by single emitter/detector combinations (serial). For parallel communication the intervening optics layer, between the board and the glass substrate, is most likely to be an array of micro-lenses (diffractive or refractive) and for serial transfer a single gradient-index (GRIN) lens can be used to collimate the optical signal. The collimated signal from a transmitting optoelec-

tronic chip is coupled into the glass plate using the HOE coupler and is confined there by total internal reflection, (in some implementations mirrors are index-matched to the surfaces of the plate [JAH94b]). A similar hologram is placed at the location on the plate where the light is to be coupled into a receiving circuit and the intervening optics images the light onto the optoelectronic detector circuitry. The intervening optics layer can be incorporated into a single HOE that both collimates and redirects the signal into the plate, (as proposed by Sauer [SAU89]). The el-

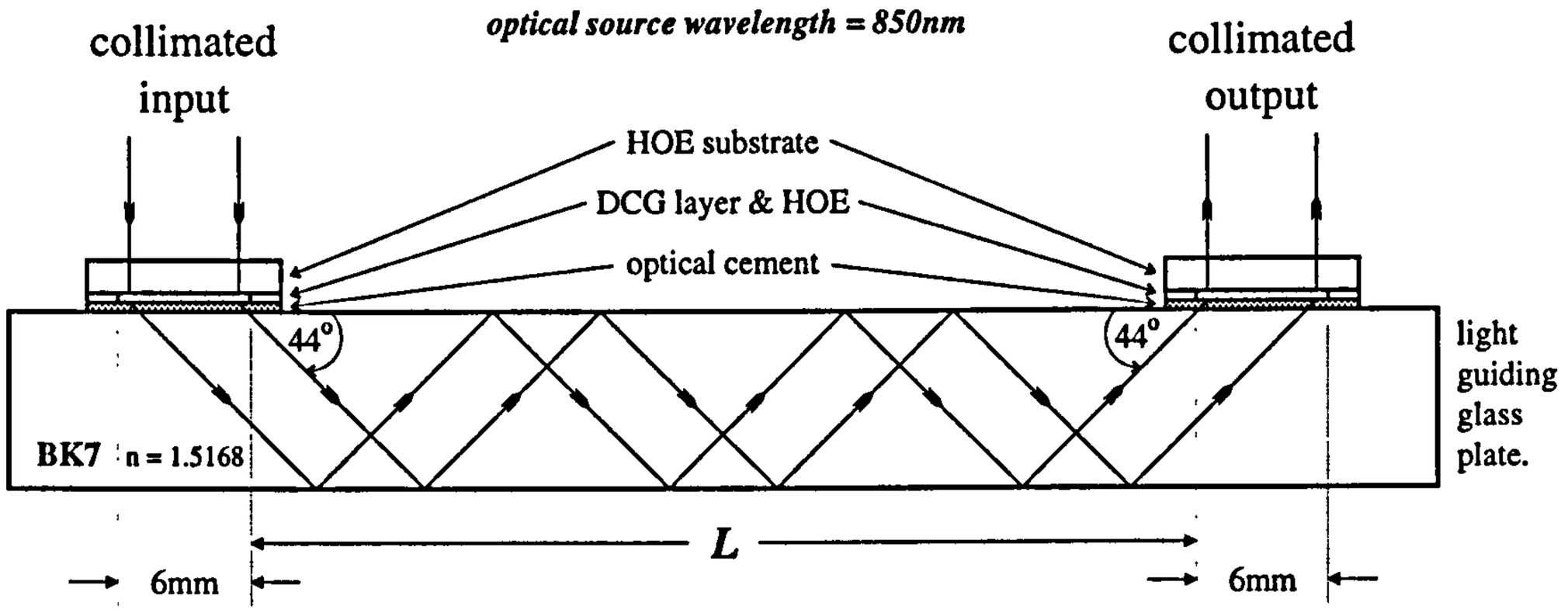


Figure 6.2: Specification and use of the guided wave coupling HOEs, (such a system can be used to achieve board-to-board interconnections in an optoelectronic system [HAU91]).

ements shown in figure 6.2 were designed and made for a demonstration system outside of the physics department (at AT&T) and are similar to those fabricated by Sauer [SAU89]. If parallel arrays of devices are to be connected in this way, the collimated beam aperture for each optical channel is typically $\sim 200\mu$m. In transferring the channels over the interconnection distance, L ($\sim$ 10cm), diffraction will cause these channels to spread, leading to crosstalk (see chapter 7 for a discussion of Gaussian beam propagation) and an overall reduction in the signal-to-noise ratio (SNR) of the communication system. This can be overcome in one of two ways:

1. Instead of having a perfectly collimated beam input to the coupling HOE, the intervening optics can impose a small amount of focal power onto the micro-beams, such that the Gaussian beam waist of the beams occurs at a distance $L/2$ within the glass plate. This leads to a slight loss of efficiency at the hologram due to a Bragg mismatch with the designed replay geometry.

2. The HOE can be designed and fabricated to include focusing power so that a collimated input will produce a Gaussian beam waist at distance $L/2$; or

3. A reflective holographic field lens (or refractive component) can be designed to be placed at a distance $L/2$ on the underside of the light guiding plate in order to compensate for any diffractive spreading [STR93] (see figure 6.3).

Each of the above have benefits and disadvantages. Option 1 is the simplest, the HOEs for such a system are described in this chapter. The optics for a board-to-board interconnect of this type, (without correction for signal beam diffraction), has been successfully demonstrated by Haumann et al. [HAU91], for 5 parallel channels, achieving data rates in excess of 500Mbits/s with an overall efficiency of 55% (over $L = 110$mm) and a SNR of more than 28dB, implying a bit error rate (BER) better than 10^{-13}. A second demonstration interconnect by the same group, Volkel et al. [VOL95], has been incorporated into a high speed telecommunication switching system. This system links six electronic boards, via 8 parallel channels, over distances of 675mm, with a SNR of 33dB and has been shown to give a BER $< 10^{-12}$ for data rates of 1.4Gbits/s. Option 2 is probably the most beneficial because, being a diffractive-pair combination, it can achieve a high efficiency and be relatively insensitive to small changes in replay angle and source wavelength [KOS93]. Option 3 has been successfully demonstrated for several parallel channels using both volume HOEs [STR91] and CGHs [JAH94b]; this solution has been shown to image, over distances between 100−200cm, close to the diffraction limit.

6.1.1 Recording and replay of coupling HOEs

The optical system used to record the HOEs (specified in figure 6.2) is shown in figure 6.4. The holograms were designed to replay at 850nm, with an angular deflection of 46° from the normal in BK7 glass, and were recorded at 514.5nm. The recording angles and index modulation required for this wavelength-shift are calculated using the techniques described in chapter 4.

The HOEs were recorded on gelatin derived from Kodak 649F plates and wash

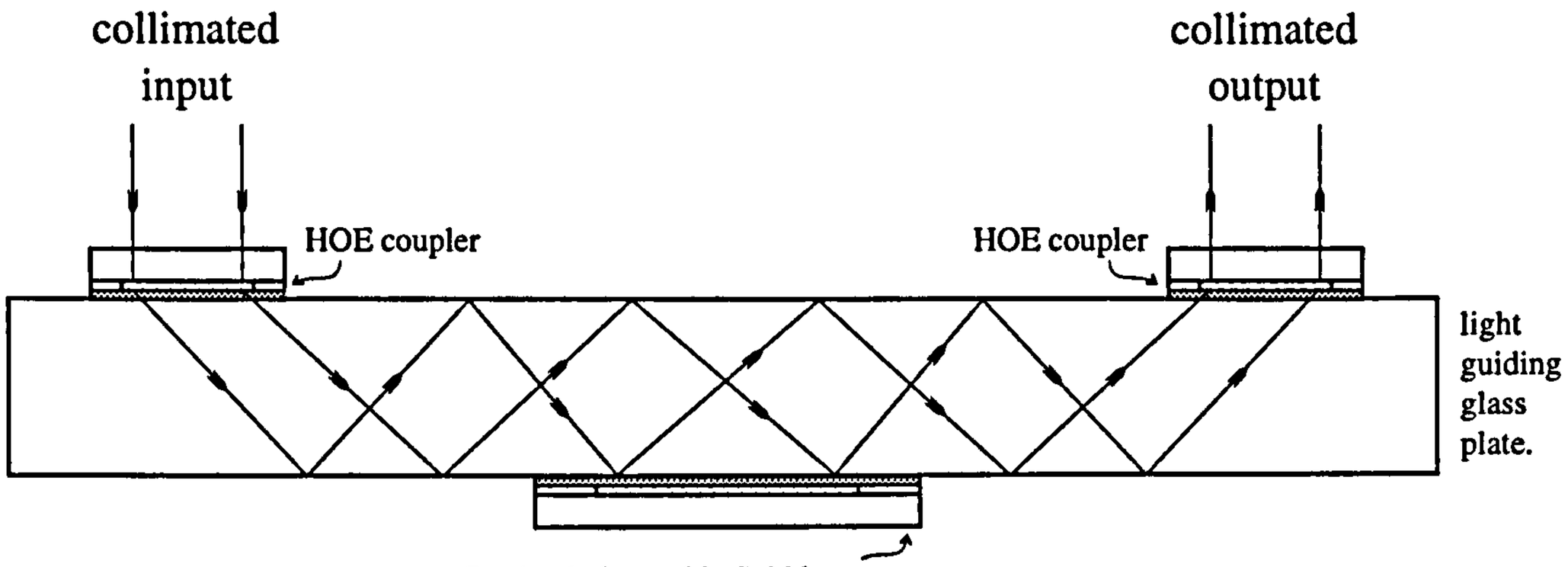


Figure 6.3: Guided wave HOE couplers used in conjunction with a reflective holographic field lens to compensate for diffraction effects.

processed. It can be seen from figure 6.4 that the apertures were not imaged onto the holographic plate because only a single large area HOE was required (6mm × 6mm), not a tightly packed array of micro-elements. Recording angles of 14.3° and 65.1°, combined with an exposure that gave an index modulation of $\Delta n = 0.0265$ at 514.5nm, resulted in a HOE with an efficiency, $\eta > 90\%$ when replayed in the configuration shown in figure 6.2.

A recording angle of 65.1° with respect to the normal of the holographic plate led to a lot of the light in this recording beam being reflected from the gelatin surface. The power in each arm of the recording system was balanced to achieve a recording beam ratio of unity (utilising all of the available index modulation). This was done by calculating the Fresnel reflection coefficients for both recording beam angles [HEC87, page 100] assuming that DCG sensitised with a 5% ADC solution had a refractive index of 1.6, (see section 3.8.2).

6.1.2 Planar coupling element conclusions

It is quite straightforward to record such elements as volume holograms with a high coupling efficiency (in excess of 90%). They have been demonstrated by several authors to be a useful way of providing chip-to-chip and board-to-board communication for optoelectronic circuits [HAU91, KOS93, STR93]. Planar interconnects can also be fabricated as CGH diffractive elements [JAH94b, SAA95], packaging of

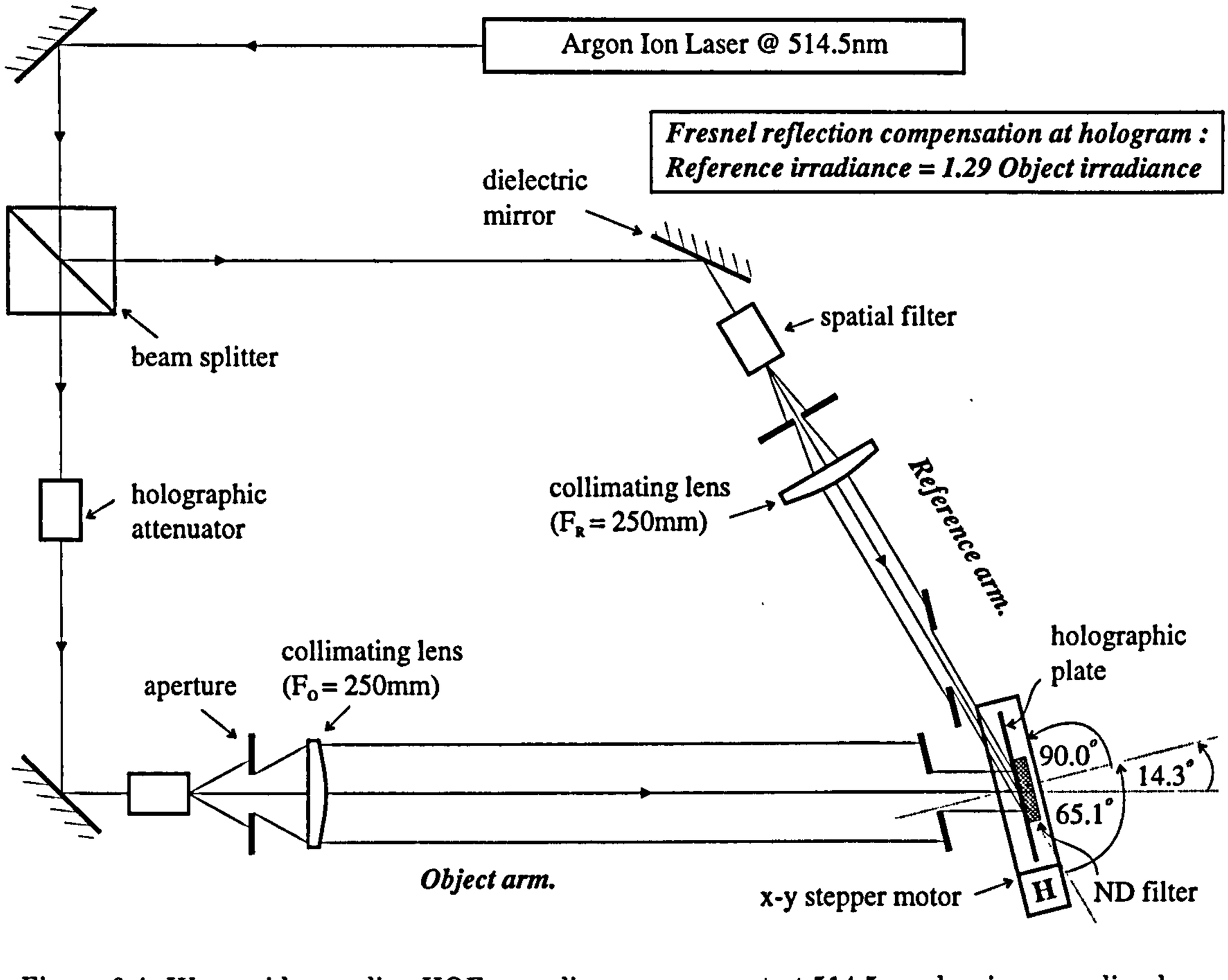


Figure 6.4: Waveguide coupling HOE recording arrangement at 514.5nm showing recording beam angles and Fresnel reflection coefficient compensation factor to achieve a beam ratio of unity.

such interconnects with optoelectronic sources, modulators and detectors has been studied in detail [ACK94, JAH94b] and they have been demonstrated as a method of optical clock distribution [WAL92].

6.2 Nearest neighbour interconnects

A nearest neighbour interconnect (NNI) is a localised connection pattern that is useful in many image processing algorithms. The essence of a NNI is shown in figure 6.5a, where the signal in each pixel or communication channel is fanned-out in four mutually orthogonal directions to the immediate neighbours. Although this does not perform a very complicated interconnection function, several powerful image processing algorithms can be implemented when it is utilised in a simple computational system implementing only basic Boolean logic e.g. a cellular-logic

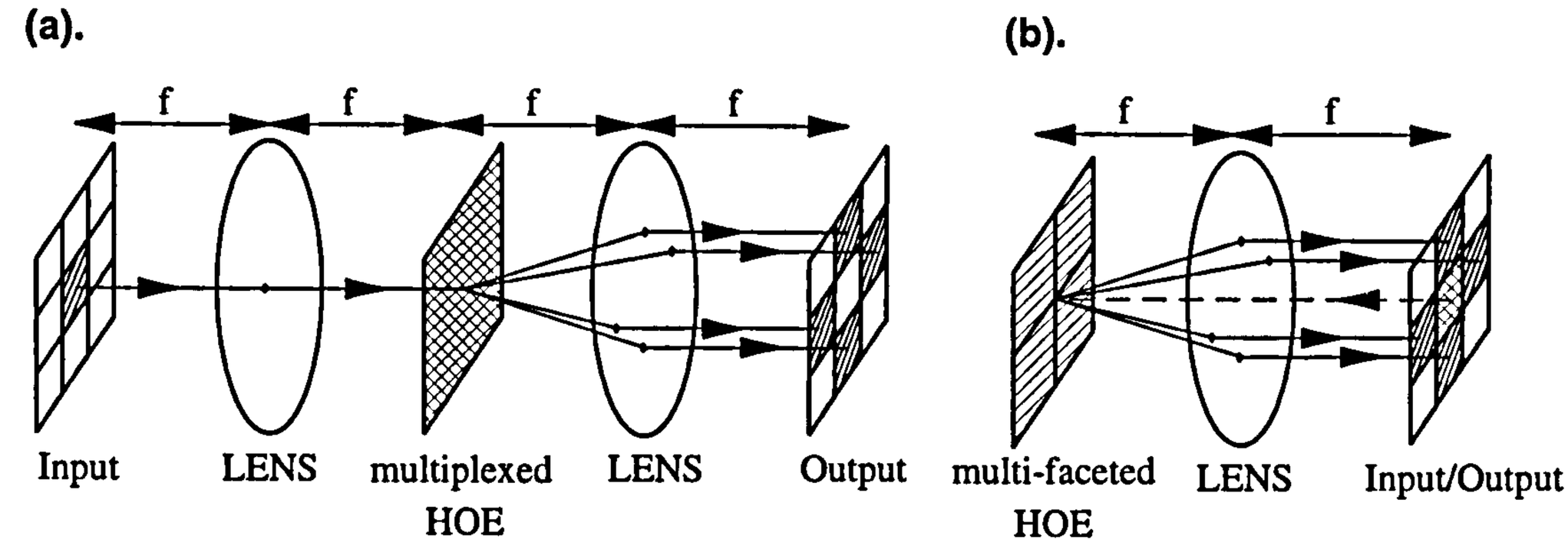


Figure 6.5: Implementations of a Fourier plane nearest neighbour interconnect in: a). Transmission [RED89b] - fan-out hologram; and b). Reflection [RES91c] - aperture divided hologram.

image processor (CLIP) [DUF86]. In addition, it is a useful interconnect in the implementation of systolic array processors [SED88, chapter 40]; a powerful, but simple, parallel architecture for solving some fundamental mathematical problems.

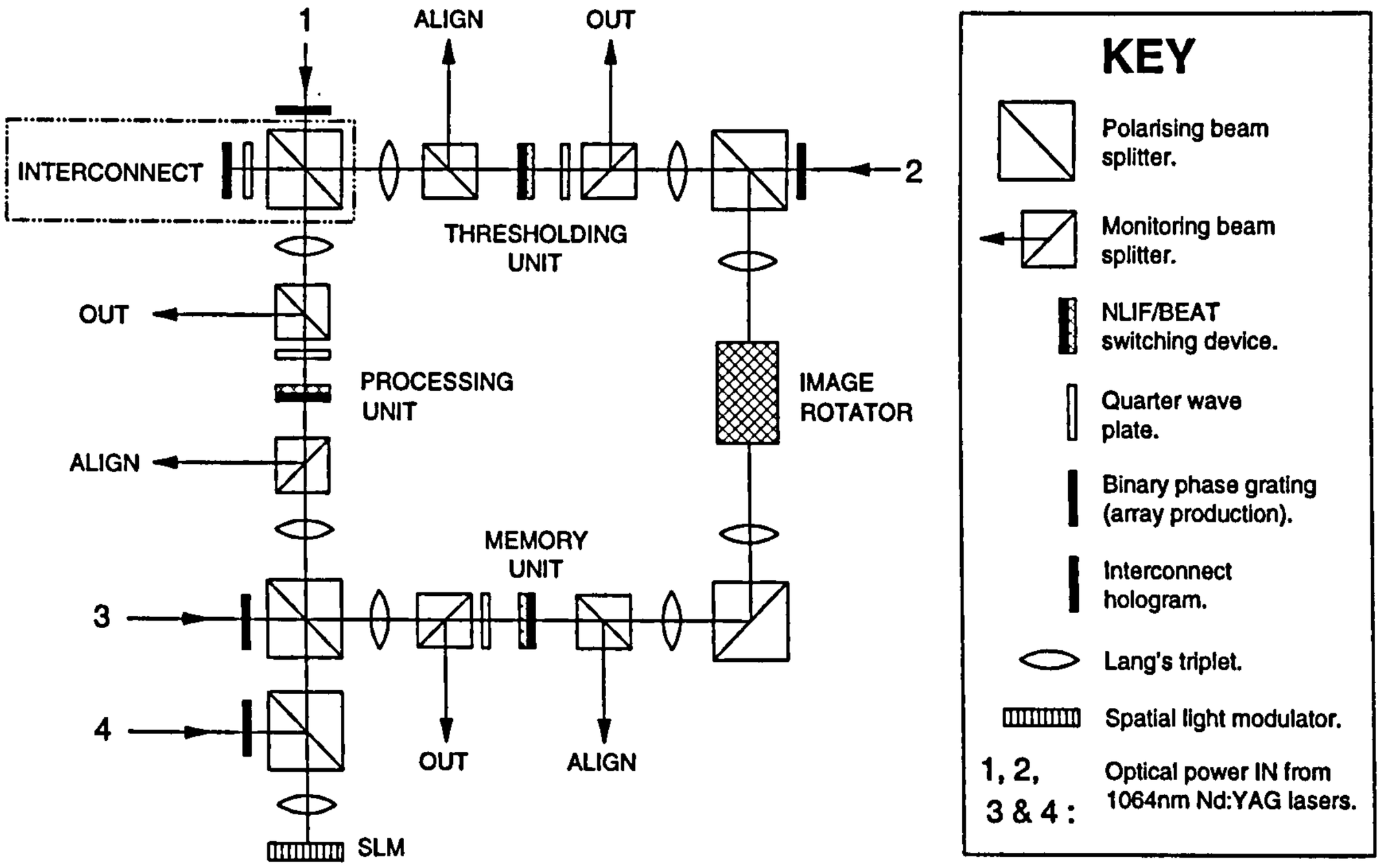


Figure 6.6: The reflection NNI used in an O-CLIP demonstration circuit [CRA91, chapter 5]. The interconnect section of the circuit shown in figure 6.12 is highlighted within the rectangle at the top left hand corner.

During the programme of optical computing research at Heriot-Watt an optical-CLIP (O-CLIP) demonstration circuit was designed and implemented [CRA91,

WAL91]. This work was undertaken to test the viability of realising an optical co-processor (residing within an electronic host computer) to perform complex image processing functions that are numerically intensive and time consuming on conventional single instruction single data-stream (SISD) electronic systems [SIE82], e.g. edge extraction (contour detection), noise removal, maze solving, target recognition and various cellular automata. The project brought together several enabling optical technologies [CRA91, chapter 5] to exploit the high degree of parallelism of optical systems and the considerable flexibility of optical interconnects; demonstrating the functional capabilities of optically interconnected logic arrays. The all-optical logic devices used were opto-thermal non-linear interference filters (NLIFs). The processor architecture was designed to manipulate regular arrays of optical signals (in accordance with the computational model of figure 1.2) and can be categorised as a single instruction multiple data-stream (SIMD) system. The optical circuitry of the 16×16 processor (256 computational elements per array) is shown in figure 6.6. The computational power of such a SIMD system is increased by a factor equal to the parallelism of the optical processor over SISD architectures and in theory it is scalable.

A nearest neighbour interconnect that was required for the demonstration O-CLIP circuit is described here. Details of the design, construction, tolerancing and results of the optical processor can be found in published work [WHE90, CRA91, WAL91]. Subsequently, further O-CLIP circuits were demonstrated using S-SEED logic arrays [TOO93, TOO94]. More complex non-local interconnections, e.g. the perfect shuffle (see chapter 7), could be used with these processors to implement more powerful algorithms such as fast sorts and discrete Fourier transforms. In principle the architecture can implement any binary image algebra algorithm.

6.2.1 Interconnect specification

The O-CLIP circuit worked with Nd:YAG laser sources that operated at 1.064μm. The 16×16 regularly spaced array elements had a channel separation of 200μm, (although the devices were not physically pixelated), and the imaging lenses were

triplets with a focal length of 94mm (see figure 6.6). These parameters dictated an interconnection angle $\delta = 0.122°$ while the 16×16 square field of the input channels (each operating with $f/6.7$ cones of light) demanded that the HOE had an aperture of at least 15mm and a Bragg acceptance that would allow a $\pm 1°$ angular detuning with no appreciable degradation in efficiency, (typically a 1% loss can be tolerated).

6.2.2 Design considerations

A spatially invariant implementation of a NNI can be realised as either a transmission [RED89b] or a reflection [RES91c] planar grating volume hologram, both operating in the Fourier plane of an optical transforming system. These options are illustrated in figure 6.5a & b respectively. Each input pixel, in the form of a focal spot from a regular array of optical signals, is transformed into a collimated beam that is incident upon the hologram. The hologram angularly fans-out the beam into four and these signals are then re-transformed into focal spots in the output plane. In the reflection system (depicted in figure 6.5b) the input and output planes are coincident because the signal reflected from the HOE is re-transformed along the input path. Each pixel in the interconnected output plane constitutes a fan-in of four separate signals from neighbouring pixels and consequently the interconnect will have inherent restrictions imposed by the constant-radiance theorem (see appendix C).

The HOE performs the same operation on each pixel in the field of the input plane and must do so with equal efficiency between pixels. When using a volume hologram in this manner account must be taken of the Bragg acceptance condition. For a predefined uniformity of power in the output plane there is a maximum acceptable input angle to the HOE and for a predefined extent of the input field this will dictate the minimum focal length of transform lenses that can be used. Transmission HOEs have a relatively narrow Bragg acceptance angle (from figure 3.5 $\sim \pm 5°$) compared to reflection HOEs [KOG69]. A reflection hologram has approximately twice the angular acceptance (measured from diffraction maximum to minimum) of a transmission hologram modulated to the same level. In addition, the peak of efficiency in the angular spectrum is very much broader for a reflection HOE; typically

four to five times at full-width half-maximum (FWHM) for the same modulation (figures 3.5 and 6.10, although not equally modulated, illustrate this point). As the level of modulation of a volume reflection grating is increased (reducing the effective optical thickness for the diffracted beam) the angular bandwidth becomes very much broader whereas the reverse is true for a transmission grating. A direct consequence of the larger bandwidth for a reflection hologram is that the overall length of the optical system performing the interconnect can be reduced by using shorter focal length lenses.

The requirement that the HOE performs the same operation on each pixel, regardless of location in the input field, means that there must be a linear relationship between the incident and diffracted angle; moving a point by a distance x in the input plane must produce exactly the same translation in the interconnected signal. This does not generally follow for diffracted light. Consider the grating equation (discussed in chapter 4):

$$\sin\theta_0 - \sin\theta_1 = \frac{\lambda}{\Lambda}\,, \tag{6.1}$$

where θ_0 is the incident angle and θ_1 is the interconnected or diffracted angle. Taking the derivative of the grating equation with respect to the incident angle gives

$$\frac{d\theta_1}{d\theta_0} = \frac{\cos\theta_0}{\cos\theta_1} \tag{6.2}$$

This can be used in the small angle regime to predict the way in which the output angle changes due to a small variation in the input angle. The only conditions that give $d\theta_0 = d\theta_1$, are: $\theta_0 = -\theta_1$ and $\theta_0 \sim \theta_1$, these represent a pure transmission and a pure reflection grating respectively.

There are two possible approaches to implementing the fan-out operation that forms the basis of a NNI:

1. **A multiplexed HOE**, constituting four gratings recorded in the same volume, each with a different interconnection angle, δ, in either the x or $y-$plane (as illustrated in figure 6.5a); or

2. **A multi-faceted HOE**, implying aperture division of the hologram into four

areas, each interconnecting through a different δ (as illustrated in figure 6.5b).

It can be concluded that the following interdependent considerations must be taken into account when designing a volume NNI hologram:

- Efficiency, η, and uniformity error, ΔU,
- Angular bandwidth, $\Delta\theta_B$,
- Hologram type: fan-out or aperture divided,
- Available refractive index modulation, Δn,
- Overall length of the optical system,
- Replay wavelength; and
- Circuit compatible geometry.

The following discussion shows that there is a clear advantage in using a reflection over a transmission hologram based upon the above considerations:

Transmission

A transmission interconnect has been fabricated by Redmond [RED89a, pages 172-176]. This fan-out hologram was designed to interconnect an array of optical channels with a separation of 2mm at a replay wavelength of 514.5nm, (the same as recording). As discussed in chapter 5 there are significant efficiency and uniformity trade-offs with transmission volume fan-out holograms, this HOE was limited to $\eta = 40\%$ and $\Delta U = \pm 2\%$.

In addition, there are problems to consider when coherently fanning-in multiple beams (in transmission) onto a transmission grating (discussed in appendix C). To achieve efficient fan-in the beams must all interfere constructively at the hologram plane, at all times, otherwise optical power remains in the zero-order beam of the hologram and is not interconnected; this leads to a degradation in the desired replay fidelity.

To maintain a flat efficiency response across the input field, (i.e. the largest deviation in the input angle being well within the angular acceptance of the grating, $\Delta\theta_B \sim \pm 4.5°$), a $4f$ imaging system with lenses of focal length 300mm was used. This made the entire system 1.2m in length. The linear variation of input with output required a pure transmission grating geometry in which the grating fringes were perpendicular to the surface of the HOE. A Bragg angle, $\theta_B = 30°$ was required to obtain the index modulation, necessary for the achievable η and ΔU, in a 14μm gelatin layer. This led to the optical system being 'dog-legged', turning through an interconnection angle of $2\theta_B = 60°$. The whole $4f$ system was both off-axis and very long, making it difficult to incorporate into the opto-mechanics of interconnect circuitry. In addition, the long focal length implied a fan-out angle, $\delta = 0.382°$ and a hologram diameter of 10mm. Therefore, the optics for the whole interconnect were designed to work at $f/30$, this is too large in comparison with an f/number ~ 5 used in many free-space communication and processing systems.

Reflection

A reflection NNI can be recorded as a multiplexed hologram with high efficiency and excellent uniformity, (as described in section 5.7), when used in the same orientation as shown in figure 6.5b (although this figure shows a multi-faceted HOE). However, there is a problem to consider regarding the available index modulation in DCG when making an HOE to replay at 1.064μm. The emulsion processing involved in creating a wavelength-shifted grating that is over-modulated to achieve a high angular acceptance is different to that described in section 2.7.2; this new processing is described in section 6.2.3. The changes in bulk refractive index and gelatin thickness inherent in this processing procedure lead to a dramatic deterioration in the replay uniformity of such a multiplexed hologram. To achieve the modulation necessary for maximum efficiency and to incorporate the wavelength-shift, gelatin layers 40μm thick are required. When these layers are processed according to section 6.2.3 they have physical parameters similar to those described in section 3.8.3. An evaluation of the performance of a multiplexed HOE in such gelatin can be obtained using the extended RCWT of chapter 5. Although this theory is only rigorously

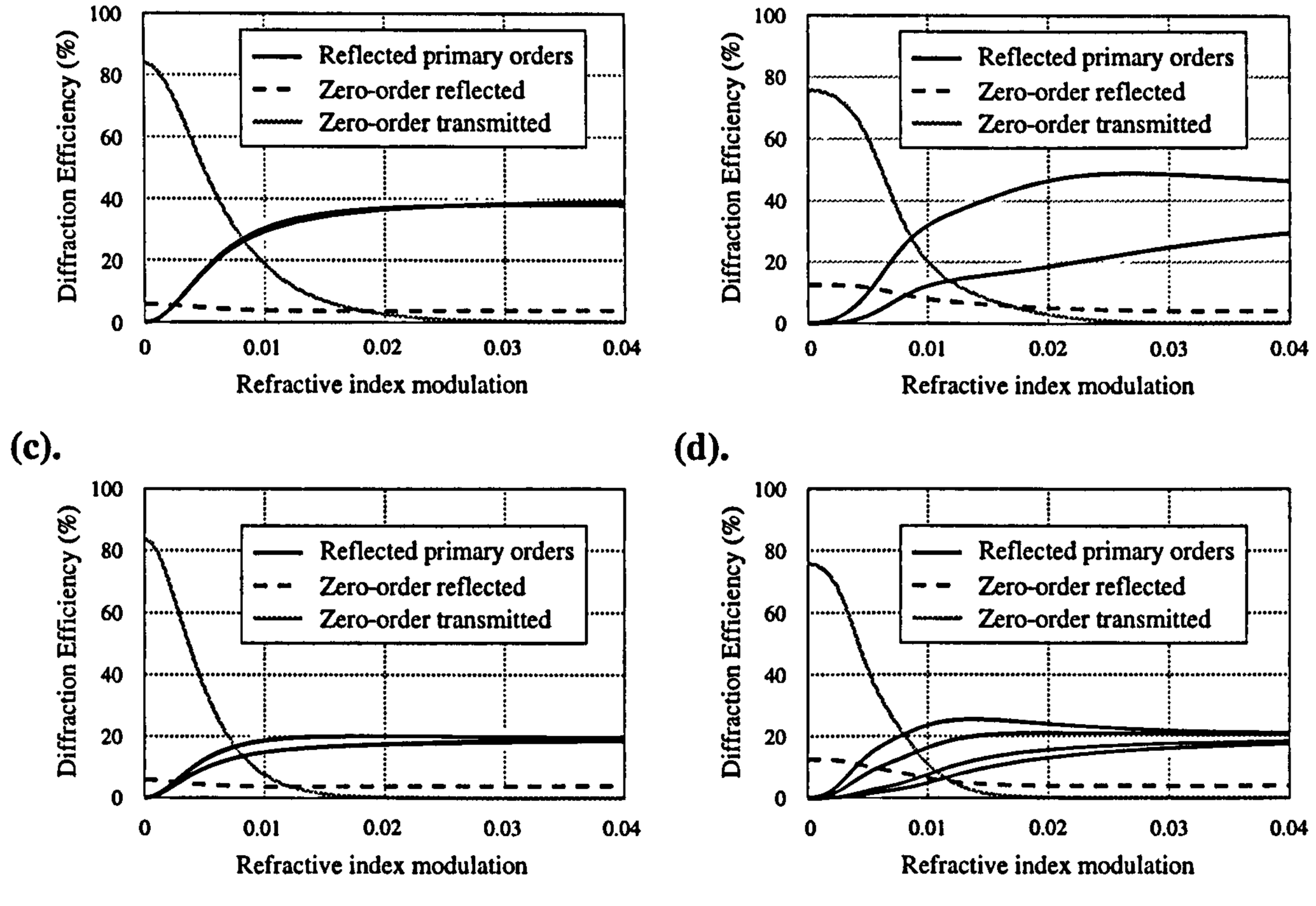


Figure 6.7: Theoretical performance of a reflection fan-out NNI (with and without DCG processing changes being taken into account), predicted using the RCWT for multiplexed gratings presented in chapter 5. The efficiency versus modulation characteristics are based upon the experimental parameters of 40μm gelatin layers used in section 3.8.3 with a fan-out angle $\delta = 0.122°$. Multiplexed gratings in 1 dimension are shown for: (a). & (b). Fan-out to two with and without processing changes respectively; and (c). & (d). Fan-out to four with and without processing changes respectively.

applicable to fan-out in one dimension it can be used to predict the behaviour of fan-out in two.

Figure 6.7 shows the expected performance of a fan-out to two and a fan-out to four reflection hologram in one dimension - fan-out in 2D should have a ΔU similar to the fan-out to two case and an η similar to the fan-out to four results. The hologram modelled with no changes in the physical parameters of the HOE between recording and replay shows both a high overall η and a good ΔU (as predicted in section 5.7). However when the changes caused by processing (as described in section 6.2.3) are accounted for, the uniformity is altered beyond acceptable limits for the application. As a consequence of this an aperture divided or multi-faceted NNI implementation is the preferred solution, when using DCG as the recording

medium.

A conformal reflection hologram can meet the specification of the NNI in a way that is convenient to the opto-mechanics of the O-CLIP circuit, except for the interconnection requirement. A linear response of output angle to input angle requires a pure reflection grating, however, for a reflection HOE to perform the interconnect, fringe planes that have a slight tilt (defined by the interconnection angle) are necessary; in this case the fringe slant angle, ϕ, is 0.061°. Experiment shows that pseudo-conformal fringes at this angle are predominantly specular in nature and hence obey a linear relationship with respect to θ_0 and θ_1. Therefore, a pseudo-conformal grating is required to implement the NNI.

The inevitable penalty for adopting an aperture division approach is that, in halving the aperture of each redirecting grating, the f/number of the interconnected signal beam is doubled, resulting in a larger diffraction limited spot size for each signal in the output plane. As a direct result of the constant-radiance theorem the irradiance of each input signal is reduced by a factor of 16 in the interconnection plane. This results from a combination of a factor of four for the 1-to-4 fan-out and a further factor of four for the doubled spot sizes and will limit the switching speed of the non-linear optical devices used in the full circuit.

6.2.3 Recording a reflection NNI

The wavelength-shifting technique described in chapter 4 was used to design the recording geometry for a reflection NNI. The recording angles of the two beams necessary to record a conformal HOE at 514.5nm for replay at 1.064μm (assuming a refractive index of 1.0 either side of the holographic plate) are such that it is impossible to couple optical power directly into the holographic medium. In order to transmit power cleanly into the DCG plate, at the correct angles and to create a noiseless HOE with no spurious gratings, right-angled prisms are required, illustrated in figure 6.8. This recording geometry has a minimum number of optical components so as to limit vibration during recording. This is important because conformal and pseudo-conformal gratings are very sensitive to having fringes washed-out due to

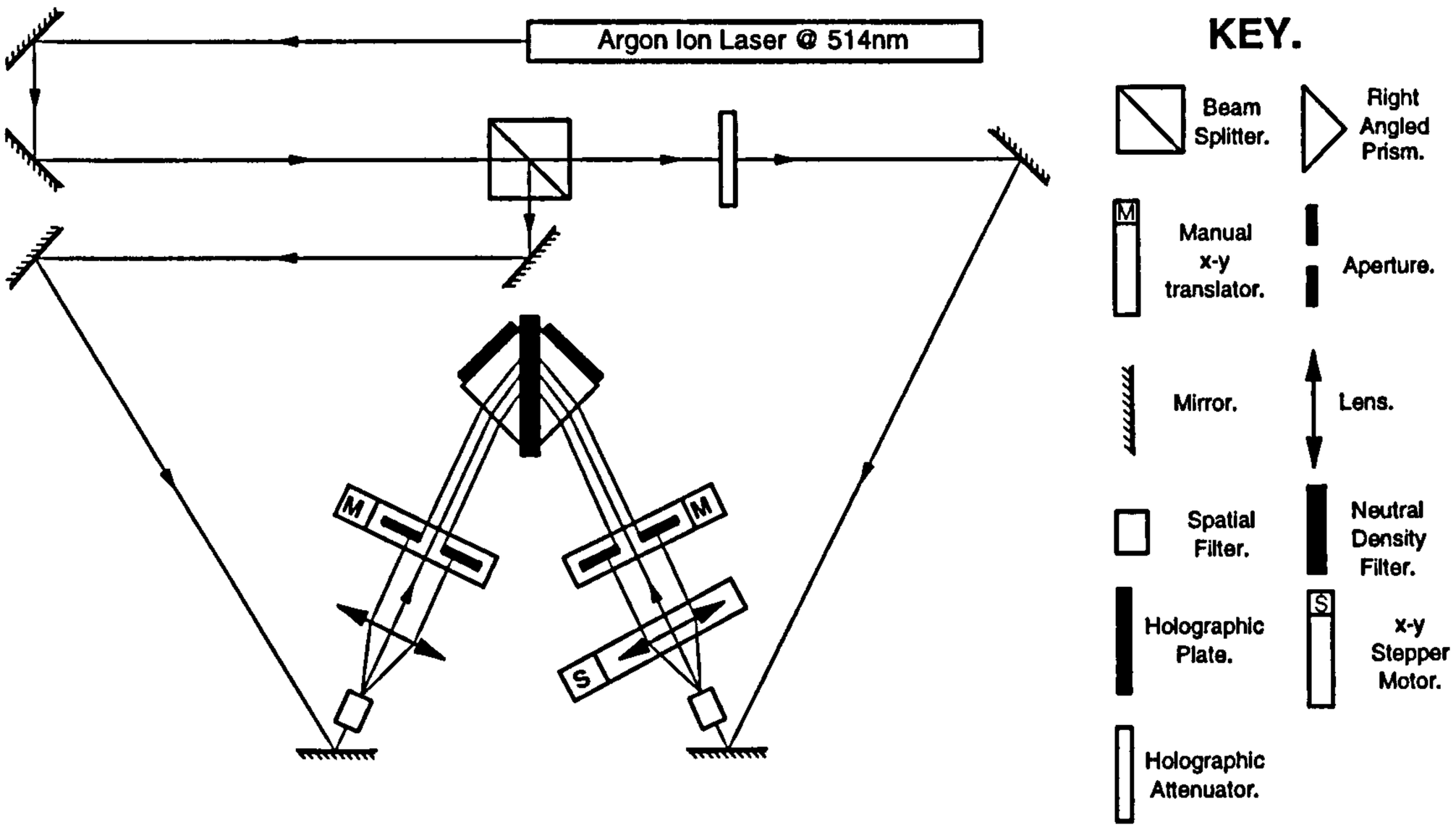


Figure 6.8: Recording geometry for the Fourier plane, faceted 1.064μm NNI reflection HOE.

opto-mechanical instabilities, (see section 2.9). The average angle of the recording beams either side of the plate inside the prisms is 61.08° to the normal, however the angle of one beam is varied in the $x-y$ plane by $\pm 0.081°$ to impose the correct grating slant angle for $\delta = 0.122°$ at 1.064μm. Imposing this slant angle is complicated by the use of right-angled prisms causing anamorphic refraction of the horizontally and vertically decentred beam. This can only be predicted accurately if the precise refractive indices of the sensitised DCG and the prisms are known. Small variations in processing will alter the thickness of the gelatin very slightly, but with a fringe slant of only 0.061° required, very small changes can significantly degrade replay accuracy. Each of the four interconnecting facets of the HOE, that redirect through angles of $\pm\delta$ in the x and $y-$plane, are recorded by decentering a collimating lens mounted on an $x-y$ stepper motor, (labeled S in figure 6.8), as proposed by Robertson et al. [ROB91a]. Changing the angle of one beam introduces the required slant angle upon the holographic fringes. The four quadrants of the NNI are recorded by manually translating the apertures (translators are labelled M) and using a different beam angle for each sector of the hologram. Figure 6.9 illustrates

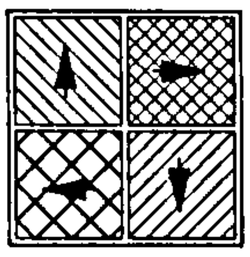


Figure 6.9: Schematic of the aperture divided reflection NNI. Each quadrant of the HOE redirects light into mutually orthogonal directions through an angle δ.

the aperture divided HOE, each quadrant interconnects through an angle δ with the sub-holograms redirecting into mutually orthogonal directions. To achieve maximum efficiency for $\pm\delta$ the recording geometry should allow for changes in both recording beam angles at once. This HOE is, however, designed to be over-modulated to such a degree that it is possible to hold one beam at the average recording angle required for a conformal grating and suffer no noticeable efficiency degradation. The high index modulation processing procedure is described in table 6.1 and details of the physical characteristics of the resulting hologram can be found in section 3.8.3, (the plate was sensitised using a 2.5% ADC solution at 20°C for 5 minutes).

High modulation processing procedure		
Lighting: dim red safelight.		
1. Water wash	2 hours	28°C
Lighting: room light.		
2. 50% Isopropanol/water	10 minutes	28°C
3. 75% Isopropanol/water	10 minutes	28°C
4. Isopropanol	10 minutes	28°C
5. Dry in circulating air	16 hours	44°C

Table 6.1: A variation on the wash-processing procedure that produces a high refractive index modulation in thick gelatin layers ($\sim 40\mu$m) and creates no significant deviation from the designed replay conditions in a wavelength-shifted grating, (water baths are used to maintain correct processing temperatures).

6.2.4 Results of the reflection NNI

The NNI reflection hologram easily meets the specification defined by the O-CLIP circuit. Figure 6.10 shows the theoretical angular spectrum of one quadrant of the HOE for the transmitted replay beam at normal incidence, and the interconnected

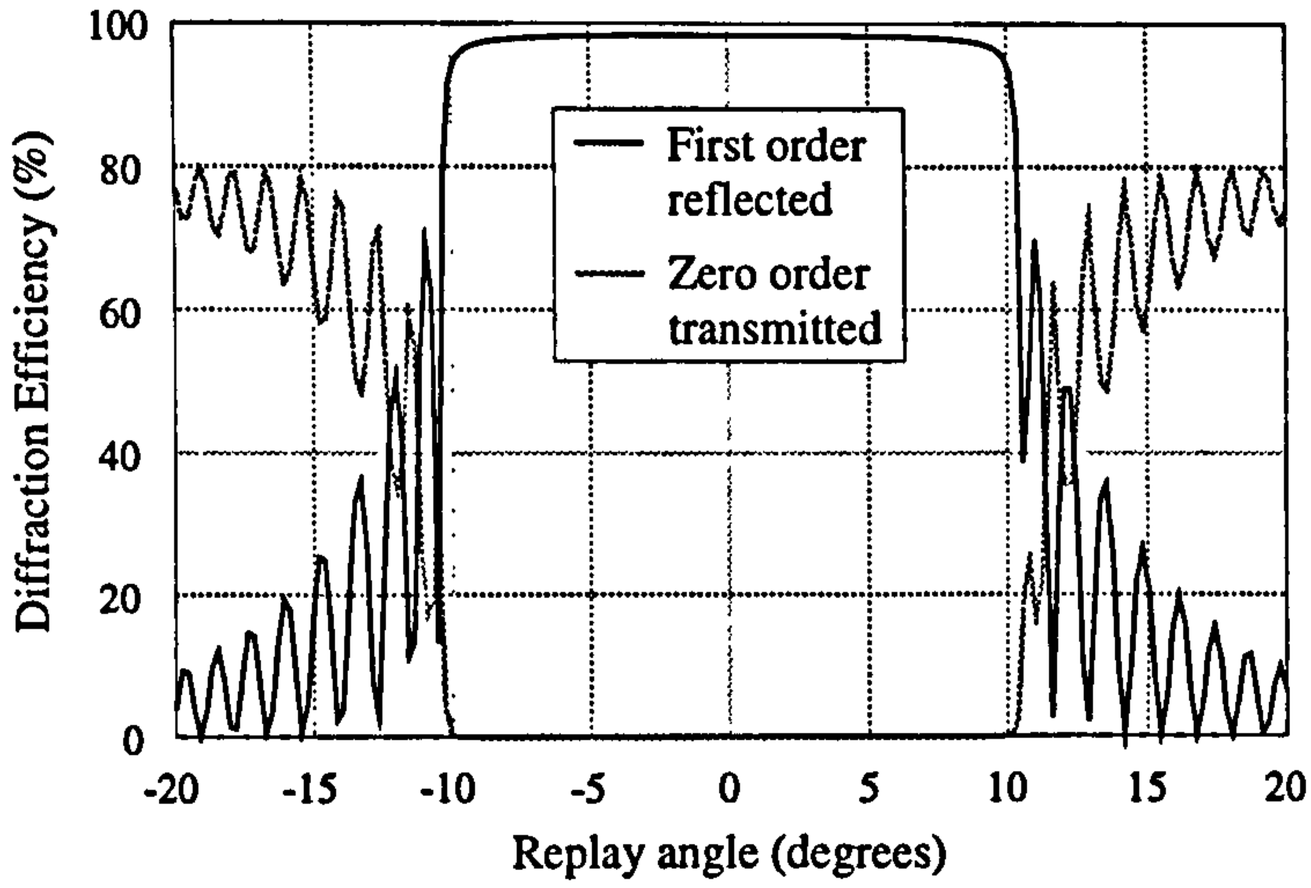


Figure 6.10: Zero- and first-order angular spectra of the multi-faceted reflection NNI predicted by the RCWT of chapter 3, (all parameters used in the model are the same as section 3.8.3), showing an angular acceptance in excess of 20°.

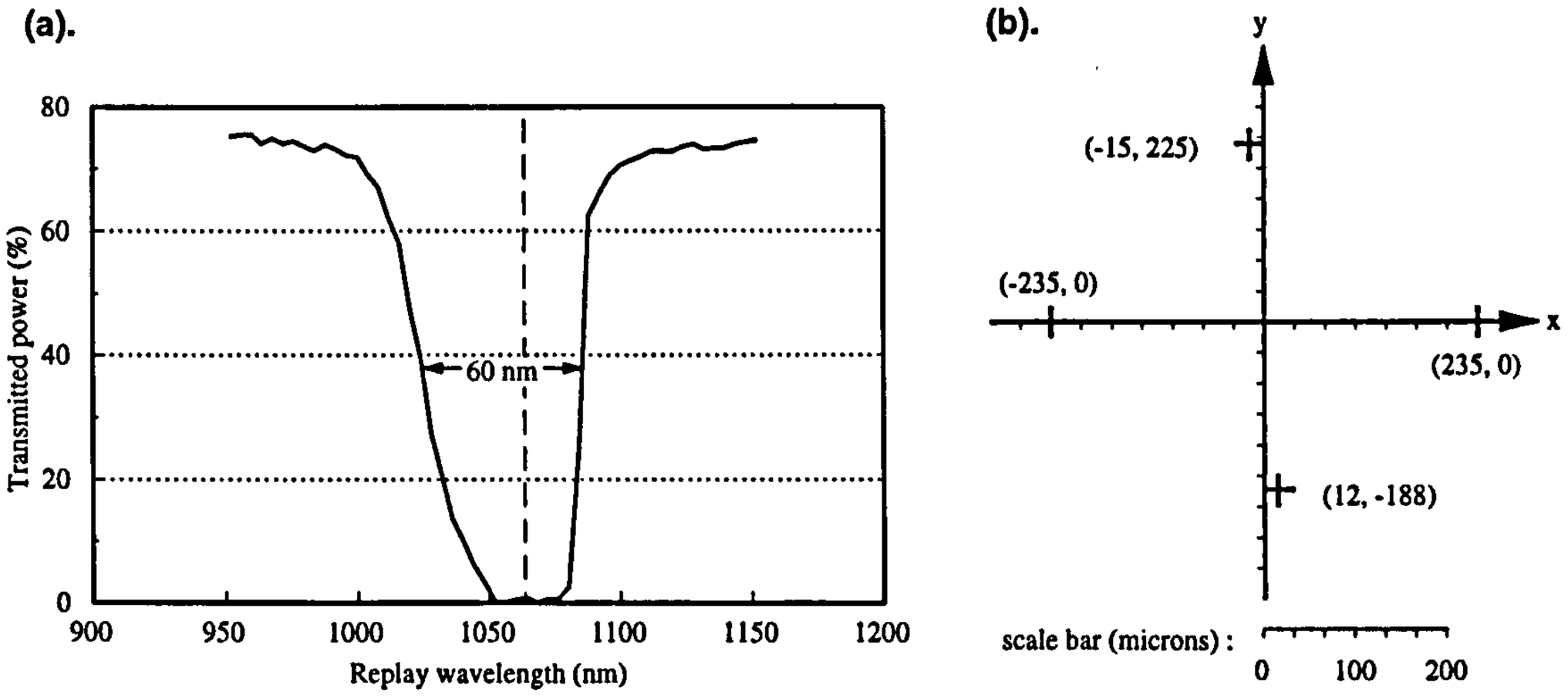


Figure 6.11: Results of the NNI reflection HOE: a). Transmission wavelength spectrum of the replay wave at normal incidence, (1064nm shown dotted); and b). Interconnected spot positions for an on-axis input to the reflection NNI, (the required co-ordinates were [±200,0] and [0,±200]).

first-order reflected wave. The gelatin absorbs about 18% of the transmitted power, however, the interconnected signal does not show any significant loss. The angular acceptance of ±10° is well beyond that required by the application (±1°). Details of the experimental replay wavelength spectrum are shown in figure 6.11a. The experimental parameters were found to be in close agreement with the theoretical predictions of the RCWT model. The optics used in conjunction with the intercon-

nect in the O-CLIP circuit are shown in figure 6.12; orthogonally polarised input and output array signals were separated using a polarising beam splitter and a quarter-wave plate (see section 7.2.3). When the hologram was replayed with a single on-axis input beam, the interconnected spot positions were as shown in figure 6.11b. The output spot sizes were 30μm (diffraction limit expected from the $f/13.4$ aperture divided interconnection optics) and they were within a spot diameter of the required interconnection positions. However, when tested with a full 16×16 array the interconnected spots were not close enough together to switch the device reliably (given the $^1/_{16}$ fan-out/fan-in loss).

The interconnected spot positional inaccuracy was due mainly to the stepper motor having a poor precision.

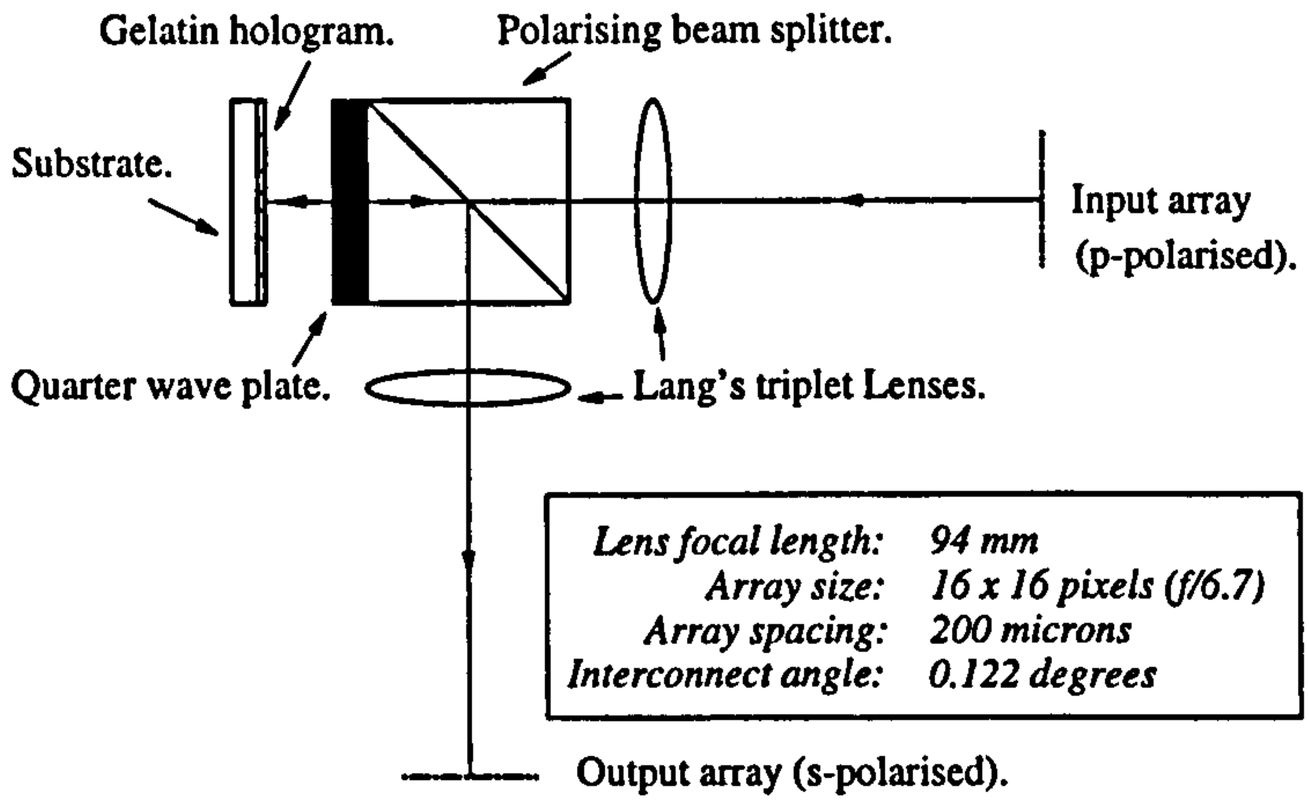


Figure 6.12: Optics used to address a reflection NNI with a regular array of signal beams and access the interconnected signal array, (input and output arrays are orthogonally polarised).

6.2.5 NNI conclusions

The interconnect could have been made to work within the strict requirements of the application given micro-stepping stages with sufficient precision. However, the multiplexed grating approach would be a preferable solution, provided a medium with sufficient dynamic range and which does not change thickness during processing is used, e.g. DuPont photopolymer.

6.3 Optical crossbar switches

The crossbar is discussed here as an example of how lenslet arrays with common Bragg angles can be used in the implementation of spatially invariant interconnection schemes. Other examples of the use of lenslet arrays in SII systems are discussed by Bird et al. [BIR91], McCormick et al. [MCC91], Streibl et al. [STR91] and Jahns et al. [JAH94a].

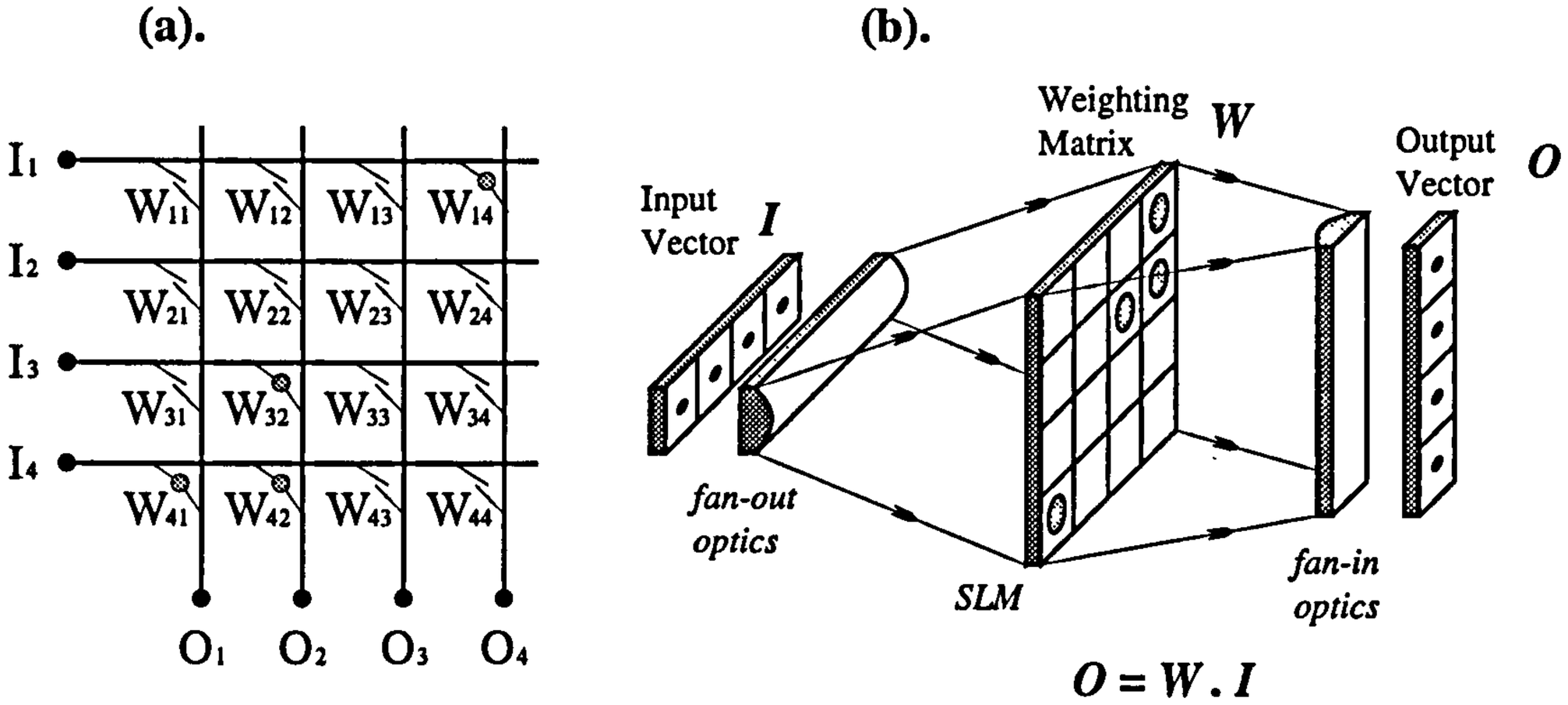


Figure 6.13: The vector-matrix multiplier crossbar switch: a). Schematic; and b). Optical implementation showing the mapping of matrix weightings with the schematic, (a 4×4 switch is shown - after McAuley [MCA86]).

The crossbar is a switching network that allows an arbitrary one-to-one interconnection of N inputs to N outputs, without blocking. A more *generalised* definition allows the number of input and output channels to be different, some inputs to have no output connection and others to have two or more [SAW87]. Figure 6.13a shows a schematic representation of a generalised crossbar with an equal number of inputs and outputs. The network is essentially a regular grid of switches, W_{mn}, between the cross-points of inputs I_m and outputs O_n; closing switch W_{mn} connects input I_m to output O_n. The inputs and outputs can be viewed as a linear vector and the grid of switches as a binary weighting or interconnection matrix. Hence the entire switch can be viewed as a vector-matrix multiplier:

$$\mathbf{O} = \mathbf{W.I}\text{; or} \tag{6.3}$$

$$O_n = \sum_{m,n=1}^{N} W_{mn} I_m ,$$

where the weights W_{mn} are 0 for switch open and 1 for switch closed.

Five distinct connection conditions are facilitated by the crossbar (see figure 6.13a):

- **one-to-one:** I_1 is connected to O_4,
- **multi-casting:** I_4 is connected to O_1 & O_2,
- **narrow-casting:** O_2 is connected to I_3 & I_4, (this requires buffering to avoid contention),
- **broadcasting:** if W_{11}, W_{12}, W_{13} and W_{14} were all closed, I_1 would be simultaneously connected to *all* outputs; and
- **no connection:** I_2 is unconnected.

The crossbar is consequently both *reconfigurable* (any combination of connections can be realised by simply opening and closing switches) and *non-blocking* (new connections can be established independently of existing ones). These key properties make it a valuable interconnect for use in parallel processing and communication architectures [GOO78, PSA84, MCA86, SAW87].

The generalised crossbar is a single-stage network (as opposed to the multi-stage interconnects discussed in chapter 7). In such an N input/output single-stage network there is an inherent fan-out to N and a subsequent fan-in to N. This leads to a weighting matrix with N^2 switching points. As the number, N, of interconnected channels (or processing elements) grows, the hardware costs of implementing such a switch electronically can become prohibitive. Allied to this cost is the associated connectivity limitations of 2D VLSI technology, as opposed to the extra freedom and bandwidth facilitated by 3D free-space optical interconnects. It is intuitively beneficial to exploit the inherent parallelism of optical interconnects to implement large crossbar networks. Sawchuk et al. [SAW87] discuss the various design issues in both electronic and optoelectronic implementations and Kirk et al. [KIR91] present

a detailed analysis of the factors limiting the scalability of free-space optical crossbar design.

The system discussed here was a prototype, used to test the viability of the diffractive and refractive optics required by an optical implementation of the vector-matrix crossbar. The lessons learnt from this were applied to the building of a much larger, matrix-matrix crossbar under the Optically Connected Parallel Machines (OCPM) project [BAR95].

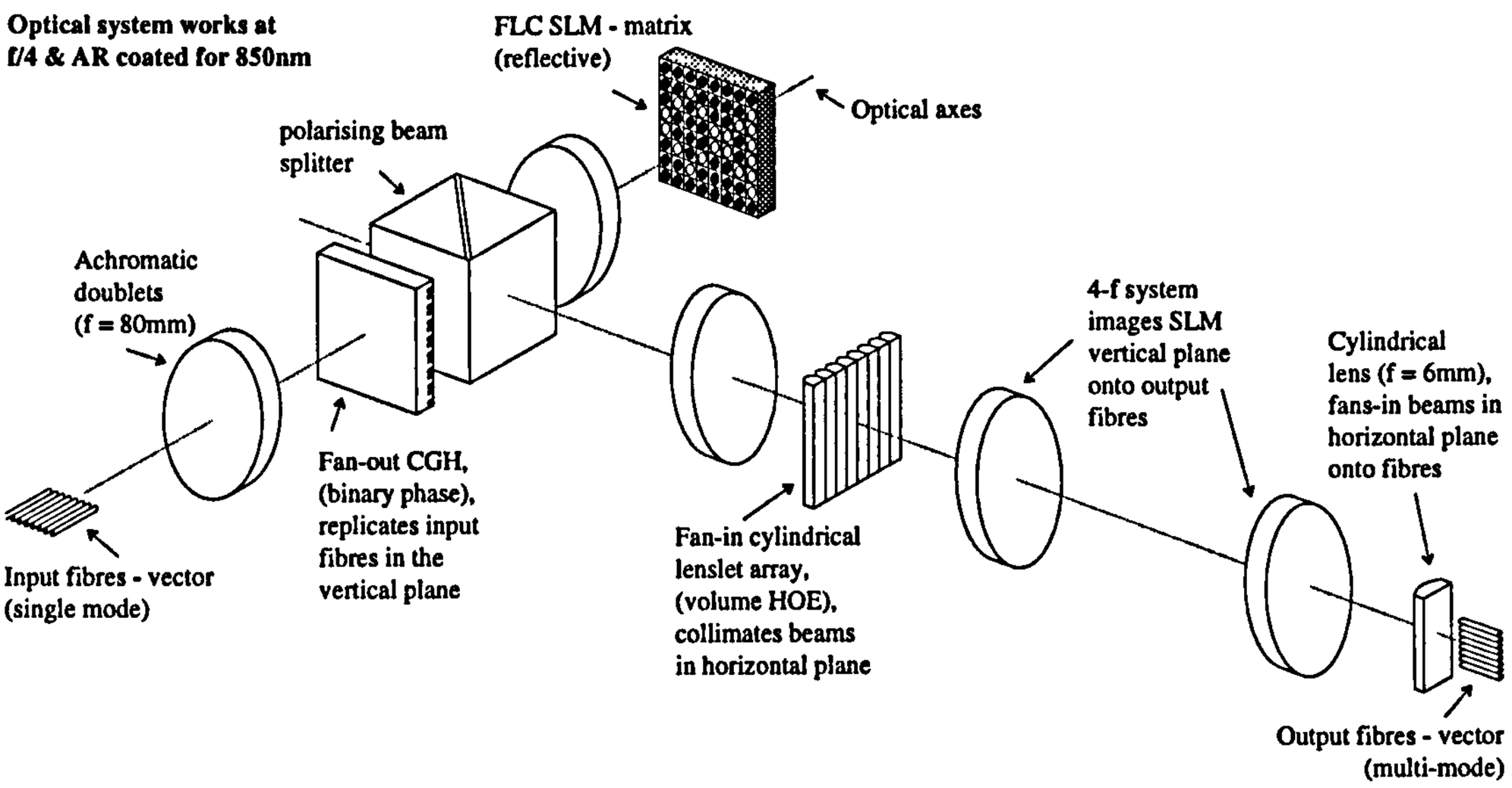


Figure 6.14: 3D representation of the vector-matrix multiplier optical crossbar switch, (an 8×8 switching system is illustrated, not to scale).

6.3.1 Vector-matrix approach

There are several ways of implementing a crossbar using free-space optics [BAR95, chapter 2], however, the most common is *shadow-masking*. Figure 6.13b shows an idealised shadow-masking vector-matrix crossbar in which cylindrical optics are used to fan-out the signal from the input vector (modulated linear array of optical sources) onto the weighting matrix or masking plane. Some form of spatial light modulator (SLM), partitioned into $N \times N$ pixels, is used to select connections by blocking or allowing the transmission of light signals through the windows of the SLM. Cylindrical optics are used to fan-in the signals from the SLM onto the output

vector (linear array of detectors). This type of switch has an inherent $1/N^2$ fan-out/fan-in loss (see appendix C) making it a very inefficient system. This, however, is outweighed by the flexibility and general connectivity of a single stage network.

The actual optical system used to realise the crossbar (shown in figure 6.13) is illustrated in figure 6.14. This is a variation on the design of the Stanford vector-matrix multiplier, first proposed by Goodman et al. [GOO78]. Figure 6.15 shows the ray path followed by one of the input signals through the interconnection system onto the output.

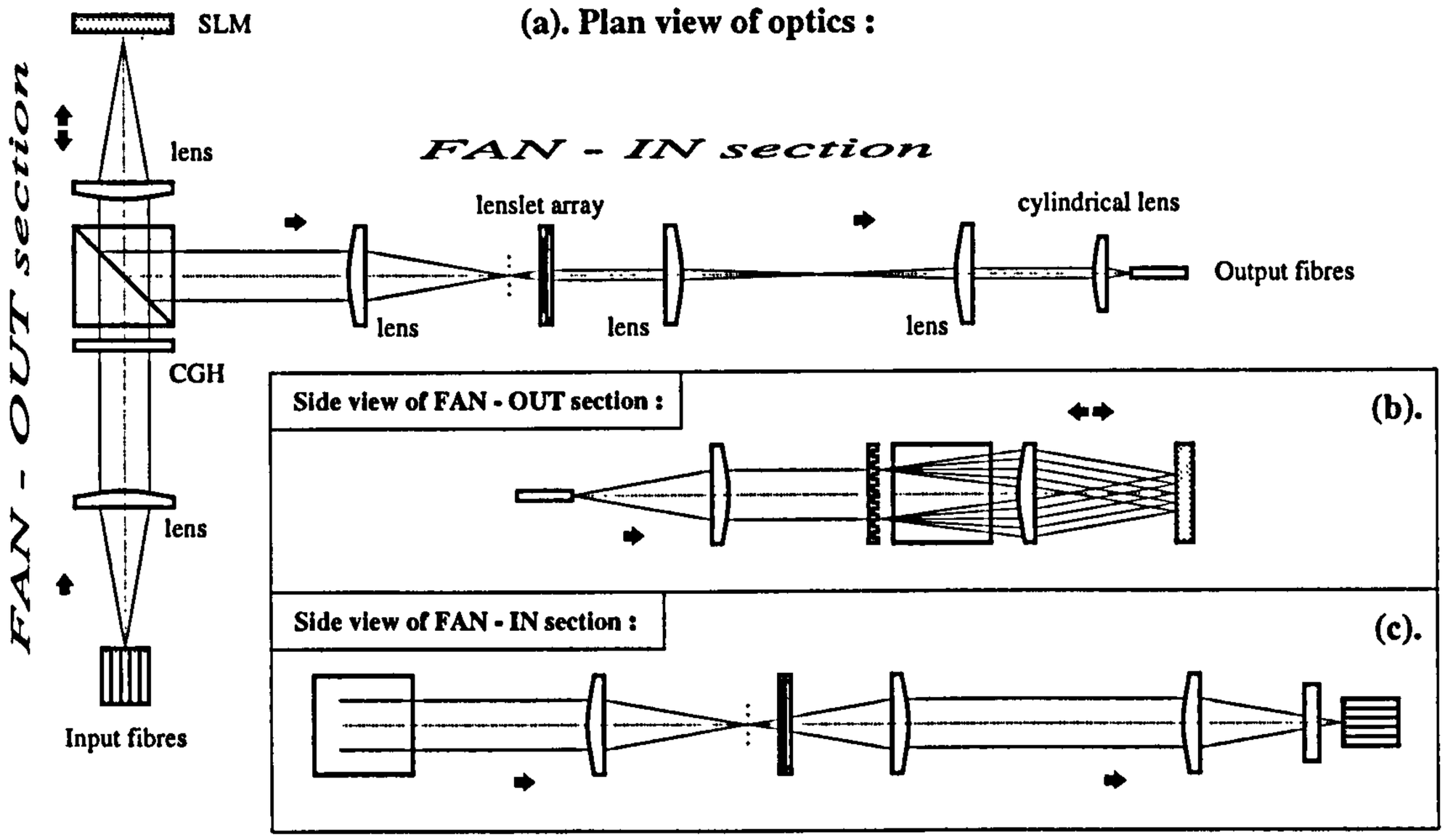


Figure 6.15: 2D representation of the vector-matrix multiplier optical crossbar switch: a). Plan view of the optics, b). Side view of the fan-out section and c). Side view of the fan-in section; showing ray paths for the central communication channel, (a 5×5 switching system is illustrated, not to scale).

6.3.2 Experimental implementation

The system was designed to be a 16×16 switch, but to be tested with a single input provided by an 850nm laser diode, pig-tailed to a polarisation maintaining fibre. All bulk optical components were off-the-shelf AR-coated for 850nm at normal incidence. It was decided that the $1/N$ fan-in loss would be avoided by using multi-mode output fibres, as discussed in section 5.1.

The one-dimensional CGH used to fan-out the input vector was a binary phase grating with a diffraction efficiency of 81% [MIL93]. The spacing of the input fibres and SLM pixels was 200μm and the N.A. of the input fibres dictated that the system should operate at $f/4$, requiring relay lenses with a focal length of 80mm. The fan-in of signals reflected by the SLM (in this case a ferro-electric liquid crystal with silicon backplane design [CRO93]) onto the output fibres was achieved using a cylindrical lenslet array and cylindrical lens combination, both were designed to operate at $f/4$.

The cylindrical lenslet array is the volume holographic component used in this interconnect. The details of its design and fabrication can be found in section 4.6.

The whole optical system was designed to work on-axis using telecentric imaging, as a consequence of this the lenslet component was a doublet structure comprising: an off-axis lenslet array and a planar grating (used to re-diffract the off-axis collimated beamlets on-axis, see figures 4.19 and 4.23). The overall diffraction efficiency of this doublet was 86%. Thus the total insertion loss in the system due to all the diffractive components was -1.57dB.

The imaging properties of the optical system were within the tolerances required for light collection by the output fibres (close to that predicted by CodeV ray tracing [ORA94]). However, aberrations at the extremes of the output field meant that custom designed lenses and a technique to reduce the size of the input field would be required for larger scale crossbars. An SLM was not available for testing the system so no exact power budget figures are available, however figures from a similar optical system [WHI93] suggest < -29dB (or 0.12% transmission) at best.

A practical large-scale optical crossbar would require an insertion loss much lower than that achieved here, since the maximum rate at which data may be transmitted through an optical switch is determined by the attenuation of the system [KIR91].

6.3.3 Matrix-matrix approach

The vector-matrix crossbar architecture is inefficient in its use of the field of the relay lenses that image the input onto the SLM and the SLM onto the output. Fan-in and fan-out occur in one dimension so that only a fraction of the available 2D aperture

is used; the field growing linearly with N, for an $N \times N$ switch. The interconnect becomes more scalable if the input and output vectors are folded into a 2D square matrix, so that the field scales with $\sqrt{N}$. This architecture is a matrix-matrix multiplier and is much more practical than the vector-matrix system.

A fully operational 16×16 matrix-matrix optical crossbar (utilising an external electronic arbitration and control system [MON93]) has since been developed under the OCPM project. This system is capable of performing with a BER of 10^{-12}, a data rate of 500Mbits/s (signal bandwidth per channel) and a total switch reconfiguration time of 140μs [WHI93, BAR95]. Optics have been demonstrated for a 64×64 crossbar with a BER down to 5×10^{-11}, a data rate of 400Mbits/s and a reconfiguration time of 20μs [BAR95, chapter 5].

6.3.4 Crossbar conclusions

A spatially invariant free-space implementation of an optical crossbar switch, utilising diffractive fan-out/fan-in components is feasible. There are two key factors limiting the scalability of the switch, both associated with the optoelectronic modulator:

- The SLM is the most lossy component in the system and consequently has the most limiting effect upon the data rate; and
- The SLM has a VLSI backplane to control the shadow-masking. Consequently it suffers from pin-out limitations similar to purely electronic chips (scaling with N^2).

State of the art electronic technology can produce a 64×64 VLSI crossbar with a data rate of 400Mbits/s and a reconfiguration time of 1.3μs [VIT93]. Although this would appear superior to the optical implementation, there are hidden overheads [BAR95, chapter 2] and scalability remains a problem. Improvements in SLM technology may soon make optoelectronic crossbars much larger than 64×64 possible, surpassing the current size and data limits of electronics.

The optical crossbar remains of interest in the fields of optical communications [BAR95], processing [PSA84] and Hopfield associative memories [HAL93].

6.4 Summary

The three examples of SIIs discussed vary widely in functionality, level of interconnection and overall complexity of the entire optical system. Consequently, it is difficult to draw general conclusions on the usefulness of such interconnects. The following four points, however, are applicable to most SII systems:

1. Where the connectivity pattern is not too complex (e.g. an optical bus), the holographic interconnect can be realised with a high efficiency ($\sim$ 90%) and good uniformity ($\sim$ 2%). This facilitates parallel optical communication systems with high data rates ($\sim$ 1Gbit/s/channel) and low BERs ($\sim 10^{-12}$).

2. Connections other than one-to-one imaging imply a degree of fan-out and fan-in. Although diffractive optical components are useful in implementing these functions, optical fan-in of many signals to a single detector/collection system carries with it inevitable limitations (see appendix C).

3. Interconnects in which diffractive components are used in a spatially invariant configuration imply the use of Fourier transforming lenses [JEN84]. The focal length of these lenses (see figures 6.5 & 6.14) make the overall size of the interconnect large in comparison with some micro-optic systems. The small Bragg angle deviation constraint required for efficient replay of volume HOEs (see section 6.2.2) also leads to the use of long focal length lenses.

4. A significant advantage of SIIs is that the holograms used to implement the interconnect are not as complex as spatially variant systems. The holograms perform the same function on all inputs and are consequently identical, making them easier to fabricate.

Spatially variant interconnects (SVIs) are more complex than SIIs, but can generally be made more compact and more flexible when implemented optically. SVIs are discussed in detail in the next chapter.

Chapter 7

Spatially Variant Interconnects

Spatially variant interconnections (as discussed in section 1.3.2) have come to form the fundamental connectivity schemes of a number of optical communication and processing architectures, e.g. [MUR88, MCC90, MCA91, CLO92, HIN94, DES95, WHE95]. Many of these utilise multi-stage interconnection topologies such as the perfect shuffle, crossover and Banyan networks. These space-variant interconnects (SVIs) have been realised in a variety of ways, using for instance: bulk optics [LOH86], micro-optics [JAH88, HUT92], planar optics [JAH90a], HOEs [HAU89, ROB90, ROB91a], diffractive components [SON93, CLO94] or combinations of these [JAH90b, SAU94, PRI95]. In general, however, the performance of these interconnects has been limited by a variety of problems, including low optical throughput, scaling restrictions and non-uniform image magnification. Consequently, methods of fabricating SVIs have become increasingly important in the field of optical interconnect technology.

To implement point-to-point spatially variant interconnects holographically requires diffractive components that both collimate and redirect/deflect light; Close [CLO75] has shown that HOEs can perform such functions. The first efficient multi-faceted space-variant volume HOEs, that permuted arrays of pixels, were demonstrated by Case et al. [CAS81] for use in wavefront and coordinate transformation systems [CAS81, BAR83]. The idea has since been used to implement many multi-stage SVI schemes, utilising one or more multi-faceted HOEs, e.g. [ROB88, HAU89,

JAH90b, ROB91a, SCH92, WAN93, SAU94, KOB95].

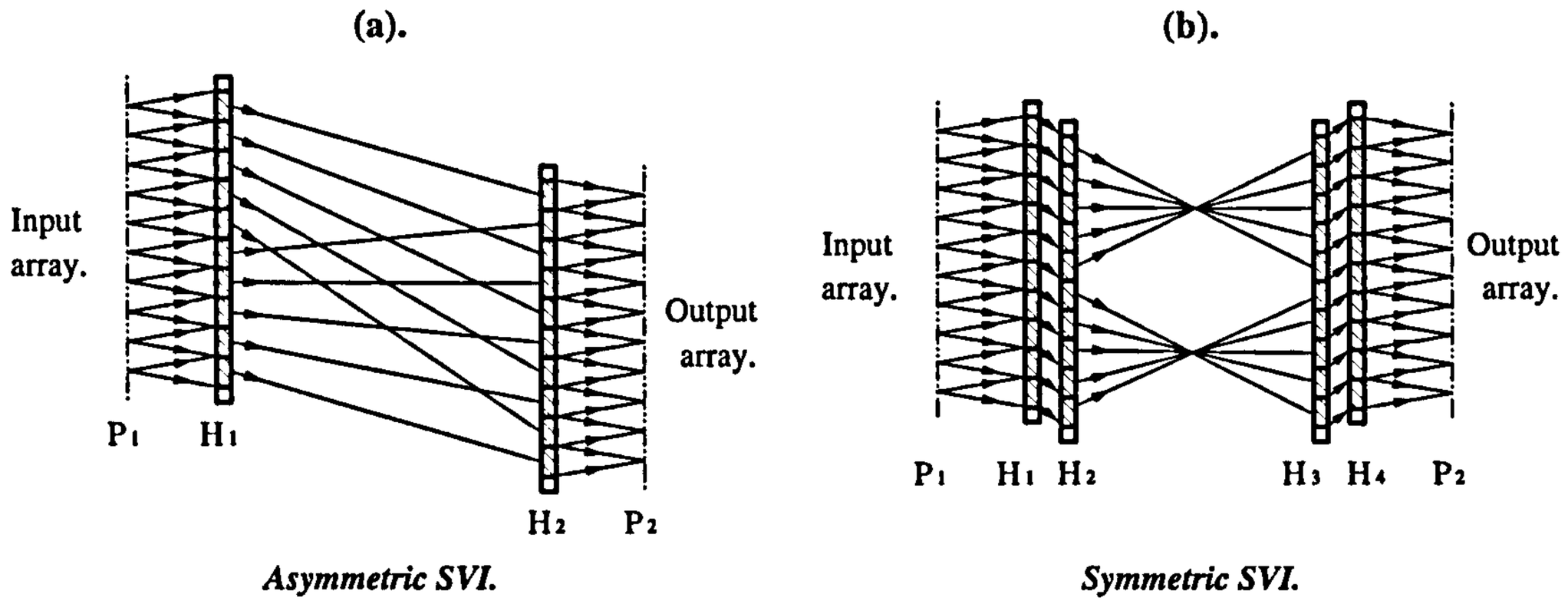


Figure 7.1: Schematic representation of asymmetric and symmetric point-to-point interconnections with: (a). A two element off-axis SVI (*perfect shuffle*) - plane P_1 is the input array, P_2 is the output (interconnected) array, H_1 is a perfect shuffling SVHOE and H_2 is an inverse perfect shuffling SVHOE; and (b). A four element on-axis SVI (first-stage of a *crossover*) - planes P_1 and P_2 are the input and output arrays, $H_{1,4}$ are identical collimating (lenslet) arrays and $H_{2,4}$ are identical, planar grating, space-variant redirection arrays.

In this chapter a method of fabricating versatile and compact, point-to-point SVIs, with volume holograms, is described. The interconnects are split into the two general categories of *off-axis* and *on-axis* SVI geometries, as shown in figures 7.1a & b respectively. In both implementations the interconnect is performed by a matched pair of multi-faceted HOEs, similar to those proposed by Haumann et al. [HAU89]. In the off-axis configuration focusing power is combined with the redirection elements while in the on-axis system focusing power is provided by a separate set of HOEs. Regular multi-stage interconnection patterns can be divided into two subgroups: symmetric and asymmetric, with the axis of symmetry lying half-way along the permutation distance and parallel to the planes of the HOEs. The matched pairs of multi-faceted HOE arrays that form the holographic SVI are consequently the same for symmetric patterns (see figure 7.1b) and inverse functions of one another for asymmetric interconnects (see figure 7.1a). The building blocks of off- and on-axis SVIs are illustrated in figure 7.2. Although it is possible to fabricate a single SVI that produces the required permutation [DAV92], all the beams emerge at the output with different angles; this is incompatible with most system requirements,

(which use telecentric imaging).

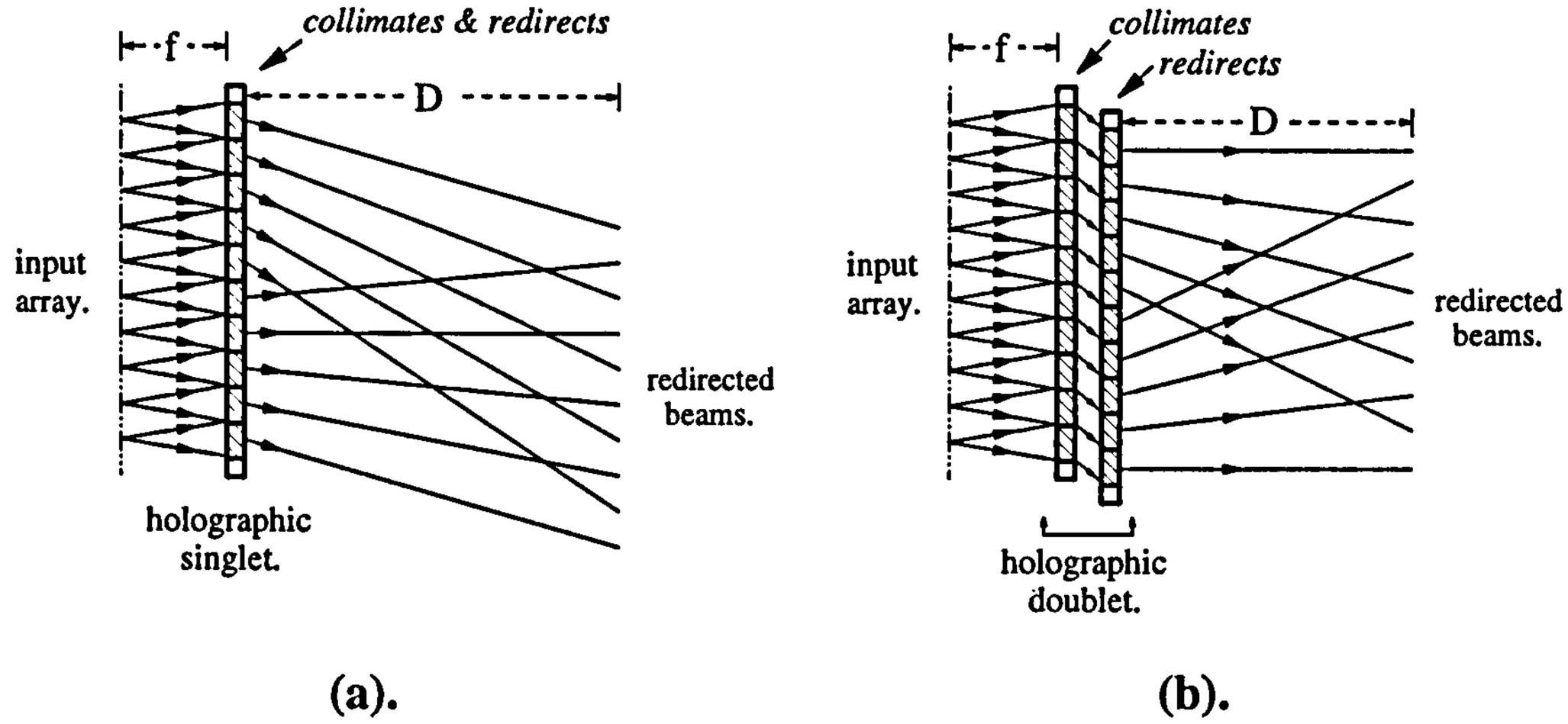


Figure 7.2: Schematic representation of: (a). The single element HOE; and (b). The doublet HOE, that form the building blocks of off- and on-axis SVIs respectively. The singlet SVHOE both collimates and redirects, whereas the doublet has separate elements to perform the collimation and redirection of the micro-beams, (both structures collimate and permute beams over distances f and D, respectively).

Experimental results for both SVI geometries are presented in this chapter. A discussion of their use with switching elements in optical communication and processing systems and the design considerations that need to be taken into account when fabricating SVIs for use in such systems, are put forward.

7.1 Off-axis SVIs

Off-axis SVIs are implemented, in this work, using the arrangement shown in figure 7.1a, (for an asymmetric interconnect). The input signals (e.g. from an array of optical switches or modulators), at plane P_1, are incident on the first space-variant HOE (SVHOE), at H_1, which consists of an array of holographic lenses. If H_1 is properly aligned one focal length from P_1, the signal beams will become collimated. As the lenses all have slightly different reconstruction (Bragg) angles, the beams cross-over to produce the required permutation at a distance D from H_1 (see figure 7.2a) where a second mirror image or inverse function SVI, H_2, is placed. This

has the effect of redirecting and concentrating the beams such that concurrent spatial focusing occurs at the output plane P_2. Although the input and output optical axes are laterally displaced all of the incident and interconnected beams travel along parallel axes, making them compatible with telecentric imaging, (common to most demonstration optical interconnect systems).

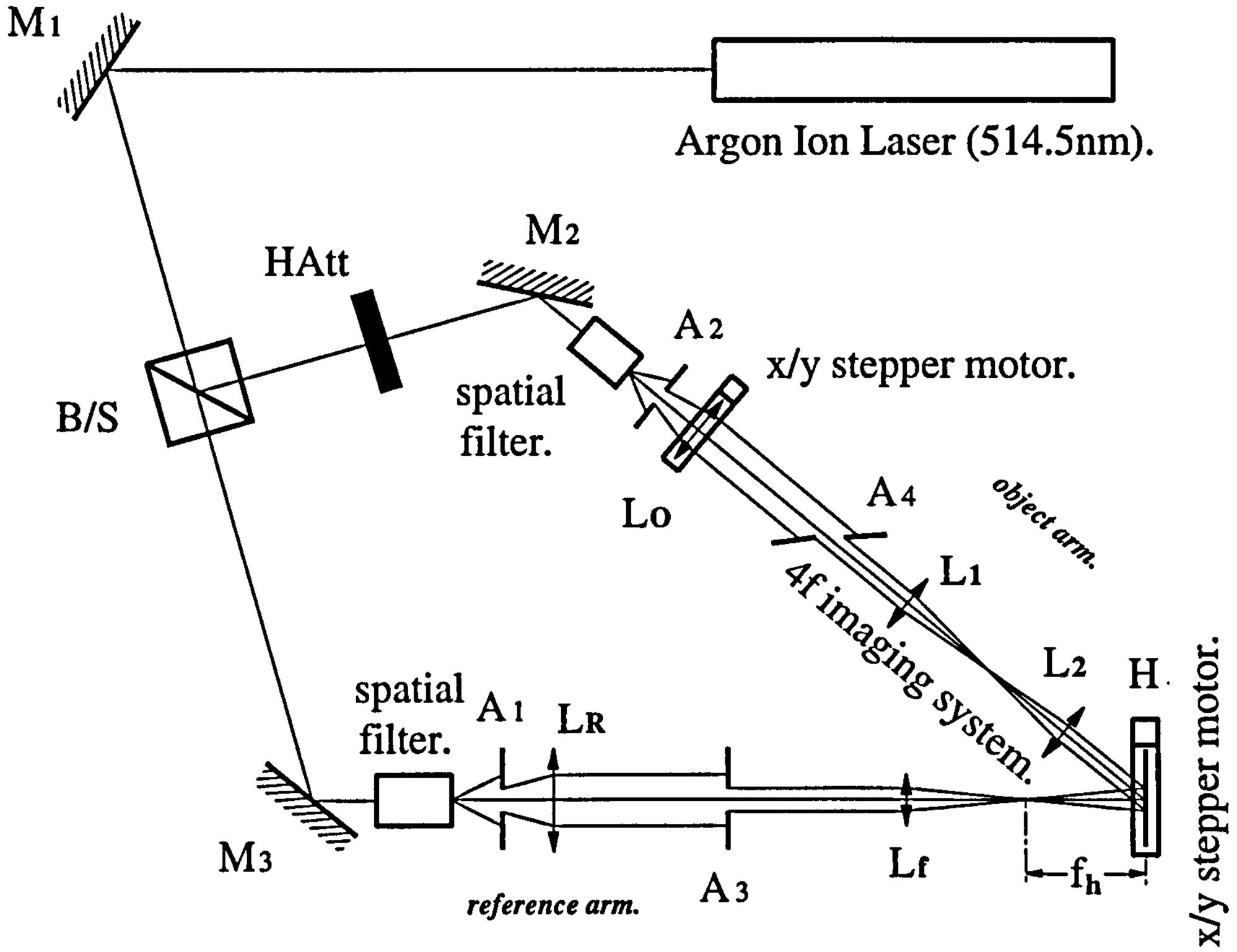


Figure 7.3: Recording geometry used to fabricate off-axis (two element) singlet SVI components: $M_{1,2,3}$ - mirrors, $L_{O,R,f,1,2}$ - lenses, $A_{1,2,3,4}$ - apertures, B/S - beam splitter; and H - holographic plane. The Bragg angle, between *object* and *reference* arms, is 30°.

7.1.1 Recording off-axis SVIs

The recording system used to fabricate singlet off-axis SVIs in DCG is shown in figure 7.3. It is essentially the same as the step and repeat exposure system used in chapter 4 to record high quality lenslet arrays. The exception being that the object arm collimating lens, L_o, is mounted on an $x - y$ stepper motor. The reference beam is derived from a spatially filtered and collimated plane wave; with an optic

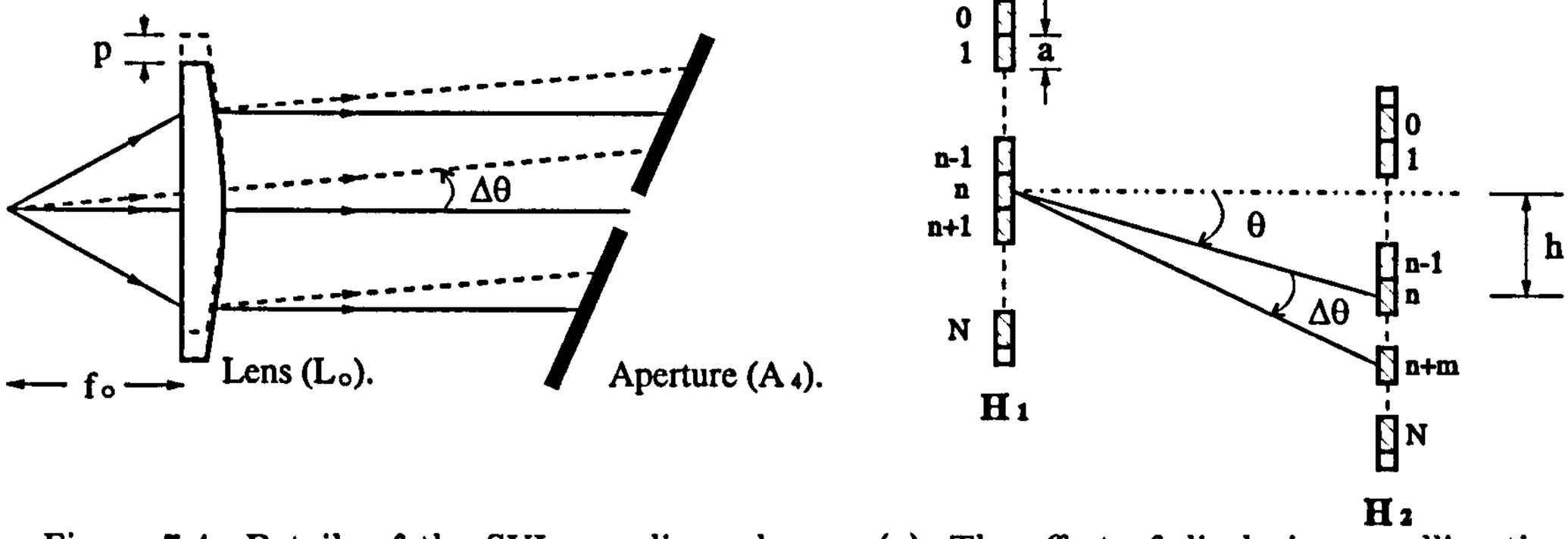


Figure 7.4: Details of the SVI recording scheme: (a). The effect of displacing a collimating lens, L_o, (also shown in figures 7.3 & 7.6b) a distance, p; and (b). Schematic of the interconnection angle, $\Delta\theta$, required to connect facet n to facet $n+m$ (between two arrays) over an interconnection distance, D, for an average replay angle, θ. On-axis SVHOEs have $\theta = 0°$.

axis perpendicular to the plane of the hologram. Lens L_f produces a point source at a distance f_h in front of the holographic plane, H, while aperture A_3 is positioned so that it is precisely imaged at H. Similarly, the object arm (at an angle θ to the reference arm) consists of a spatially filtered and collimated beam that illuminates the aperture A_4. A $4f$ lens system (L_1 & L_2) then images this aperture at the holographic plane, accurately overlapping with the object beam. Consequently, each facet of the SVI has well defined edges, allowing the maximum possible interconnect density to be achieved.

A variation, $\Delta\theta$, in the average object beam replay angle, θ, is obtained by translating the collimating lens, L_0, a distance, p, normal to the optic axis of the object beam. This causes a slight angular displacement of the collimated beam with respect to the aperture A_4, as illustrated in figure 7.4a. The angular change varies with the movement of the lens according to the relation

$$\tan(\Delta\theta) = \frac{p}{f_o}, \tag{7.1}$$

where f_o is the focal length of lens L_o. If a precise $4f$ imaging system is used, the wavefront present at A_4 will be reproduced at H. Therefore a translation of L_o will cause a change in the object beam angle and hence in the replay angle of

the reconstructed wavefront; no translation results in the HOE facet reconstructing at the standard replay angle, θ. To relate this to an actual spatial permutation of interconnected beams, consider figure 7.4b. If the two interconnecting SVHOE planes are separated by a distance D and each (square) facet has dimensions of $a \times a$, then to link facet n to facet $n+m$ requires an angular displacement of $\theta + \Delta\theta_m$. This can be derived using the relation

$$\tan(\theta + \Delta\theta_m) = \frac{h + ma}{D}, \tag{7.2}$$

where h is the lateral displacement of hologram plane H_2 with respect to H_1 and is related to the standard replay angle by $h = D \tan\theta$. Substituting equation 7.2 into equation 7.1 gives an expression for the distance, p_m, by which L_o must be translated, in order to obtain a change in replay angle $\Delta\theta_m$:

$$p_m = f_o \tan\left[M\left(\tan^{-1}\left(\tan\theta + \frac{ma}{D}\right) - \theta\right)\right], \tag{7.3}$$

where $M = f_{L2}/f_{L1}$ and $f_{L1,L2}$ are the focal lengths of the lenses in the $4f$ imaging system.

To record a SVHOE with this system, lens L_o is simply translated to give the required object beam angle and a facet is then exposed. By successively stepping the holographic plate vertically and/or horizontally and repeating the process, a complex 2D space-variant interconnect can be realised.

The advantage of this recording technique lies in the accuracy to which apertures A_3 and A_4 can be imaged and overlapped at H. Problems can, however, arise when trying to record large replay angle variations:

1. The maximum variation in the recording angle, $\Delta\theta_{max}$, that can be realised during fabrication, is limited by the numerical aperture (NA) of the recording optics; and

2. As larger angular variations are approached, aberrations become more significant.

These two key factors impose an upper limit on $\Delta\theta_{max}$ of only a few degrees. This

in turn restricts the minimum allowable hologram separation D for any particular SVI scheme.

Several other factors have to be considered in this optical recording process in order to avoid fabrication errors that can restrict an interconnects usefulness:

- A precise overlap must be maintained between the object and reference beams at H.

- Facet size and spacing must be carefully matched to the channel spacing in the system in which the SVI is to be used.

- The focal length in the recording set-up must match the NA of the SVI system.

- Highly accurate stepper motors are required to avoid positional errors due to nonlinear translation, backlash or wobble.

- The amplitude of the wave incident upon aperture A_4 (see figures 7.4a & 7.3) must be uniform in profile so that the beam-ratio (between object and reference arms) does not vary with $\Delta\theta$. This is achieved by expanding the object beam at the spatial filter so that only a uniform part of the Gaussian is used. However, this can lead to low levels of irradiance at the hologram, requiring long exposures. In certain circumstances this problem can be avoided by allowing the beam-ratio to vary and making test exposures for each $\Delta\theta$ of the interconnect, thus allowing optimum exposure for each interconnection angle to be found. If, however, the available index modulation is limited, the reduced modulation depth (caused by the varying beam-ratio) will restrict the maximum achievable efficiency.

- DCG processing changes must be compensated for (see section 2.8) so that the reconstruction angles of the SVI singlet arrays match one another and meet the design specification; and

- The recording process described works for point-to-point, two element, singlet SVIs designed to replay at the recording wavelength. If, however, a

wavelength-shift is desired the process becomes considerably more complicated, since each lenslet can have a different Bragg angle upon replay. A different recording geometry for each facet is then required and the recording techniques described in section 4.3 need to be built into the design and fabrication process. This has not been attempted since on-axis SVIs (discussed in section 7.2) provide an easier solution to this problem.

An alternative recording system that circumvents some of the NA limitations of this scheme has been used by several authors [ROB88, HAU89, KOB94]: A tilting mirror provides the varying recording angles and apertures are placed very close to the holographic plate in order to ensure that facets have well defined edges. SVIs with dimensions very similar those described here (and in section 7.2), have been fabricated with this alternative technique.

7.1.2 Replaying off-axis SVIs

The off-axis nature of these interconnects makes it difficult to define the optic axis and hence to align the two SVHOEs with respect to one another; as well as to the external optical system of which they are a part.

Each array of space-variant components has *six* degrees of freedom making alignment a considerable problem. Misalignment effects were found to limit the performance of the SVIs, leading to higher levels of crosstalk (between channels) and distorting the regularity of the output array. Thus spot size, position and power will vary across the output plane if care is not taken to align the system. Schwider et al. [SCH92] have proposed the use of mechanical distance etalons to reduce the number of degrees of freedom to *three* per array, (one rotation and two translation), in order to minimise alignment tolerances.

McCormick et al. [MCC91] have analysed the alignment errors for a pair of on-axis lenslet arrays. This analysis showed that the focal length and positioning of the facets must be tightly controlled to obtain optimum performance, especially when small channel separations are required. Consequently, the opto-mechanical design of the interconnects and packaging technology used are exceptionally important.

Apart from alignment sensitivity there are several other factors that should be considered when replaying off-axis SVIs:

- The holograms exhibit a polarisation dependence due to the different Fresnel transmission coefficients introduced by an off-axis geometry. The polarisation dependence of individual gratings, for $\Delta\theta_{max}$ of a few degrees, is negligible in comparison [RED89a, pages 91-139]. Thus for a standard replay angle of $\theta = 30°$, the transmission coefficients of the $s-$ and $p-$polarised components will be 94.2 and 97.5% respectively. On-axis systems avoid this problem since $\theta = 0°$. Kobolla et al. [KOB95] have designed interconnects to permute the beams in a glass block, index matched and cemented to the SVI pairs, in order to eliminate off-axis Fresnel losses. Where complex 2D permutations through many degrees are concerned the polarisation dependence of individual gratings becomes important and must be compensated for in order to maximise efficiency and uniformity, Kobolla et al. [KOB93] have shown how this can be done.

- The zero-order (undiffracted) light from each array of holograms either propagates out of the optical system or is easily blocked from reaching the output plane [KOB95]. Any stray light incident upon a facet of either array will, in all probability, not be Bragg matched to that volume HOE and will therefore not be diffracted as noise into the output array.

- Most demonstration optical communication circuits use in-line optics. The off-axis nature of these elements means that they are not easily incorporated into the opto-mechanics of the external system. However, this problem has been used to an advantage by Gluch et al. [GLU93] in a demonstration system where the pitch of optical channels changes between the input and output planes of the interconnect; a requirement imposed by other circuit components. Schwider et al. [SCH92] propose the use of a mirror, along the length of the base of the permutation distance, to make an in-line geometry possible with a pair of off-axis singlet elements.

- The off-axis interconnects are sensitive to chromatic changes in the replay system. Changes in the replay wavelength, $\Delta\lambda$, can dramatically alter the uniformity, size and position of the spots in the interconnected array [SCH92, SAU94]. In contrast, the close-cascaded nature of on-axis SVI doublets reduces this sensitivity [STO85, KOS89b]. Kobolla et al. [KOB95] have demonstrated a system with mixed off- and on-axis redirection elements that can tolerate a range of $\Delta\lambda = \pm 10$nm with no appreciable effect on the output.

7.1.3 Experimental results

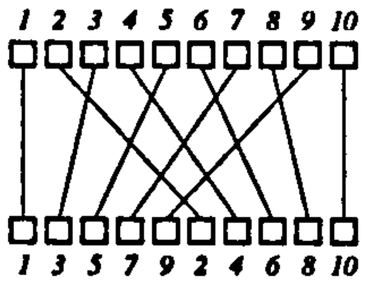
A range of prototype off-axis SVIs have been fabricated using the recording system shown in figure 7.3, including several stages of a crossover, Banyan and perfect shuffle network. All of these demonstrations were stacked arrays of 1D interconnects - this makes alignment easier and eliminates most polarisation problems when replayed with TE polarised light.

All of the holograms were recorded and replayed at 514.5nm using DCG derived from Kodak 649F plates and standard pre-processing and development procedures, (see section 2.7). In the recording of most of these interconnects $L_1 = L_2 = 100$mm.

Figure 7.5 illustrates the performance of a 10×10 facet stacked 1D inverse perfect shuffle interconnect. Figure 7.5a shows the idealised masked input to the array, produce by a loop of fine wire in the input focal plane, (labelled P_1 in figure 7.1a). Figure 7.5b shows the idealised output from the interconnect in which alternate columns of the input image are in effect split into two images of the original. Figures 7.5c & d show the resulting experimental input and output of the interconnect. In this particular case the facet size, a, was 500μm, the standard replay angle, θ, was 30°, the interconnection distance, D, was 5cm and the replay focal length, f_h, was 2cm at 514.5nm. The efficiency of each SVI singlet array was measured to be 90%, giving the overall interconnect a diffraction efficiency of 81%, (neglecting Fresnel losses).

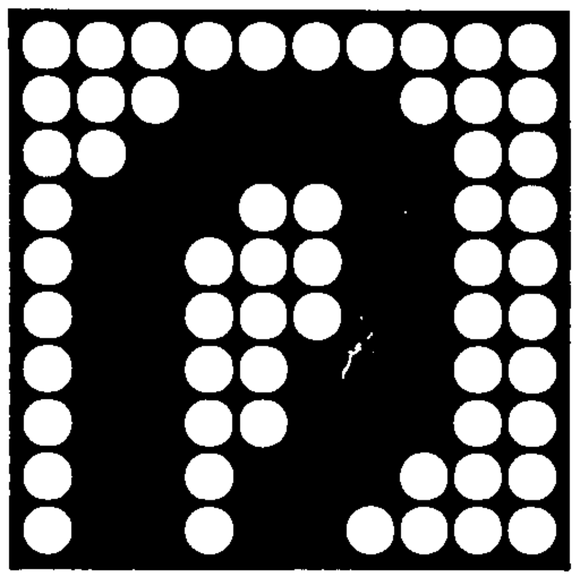
It can be seen from figure 7.5, that there is a slight positioning error in the

Schematic :



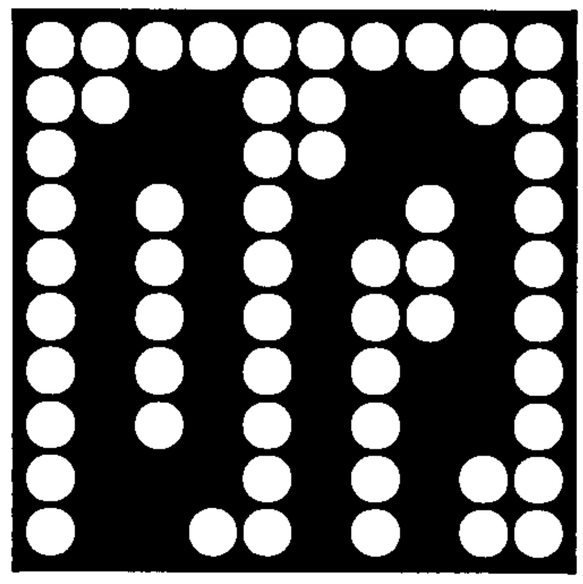
1D Inverse Perfect Shuffle schematic for 10 facets.

(a).



Masked input to stacked 1D Inverse Shuffle (loop of fine wire accross the input field).

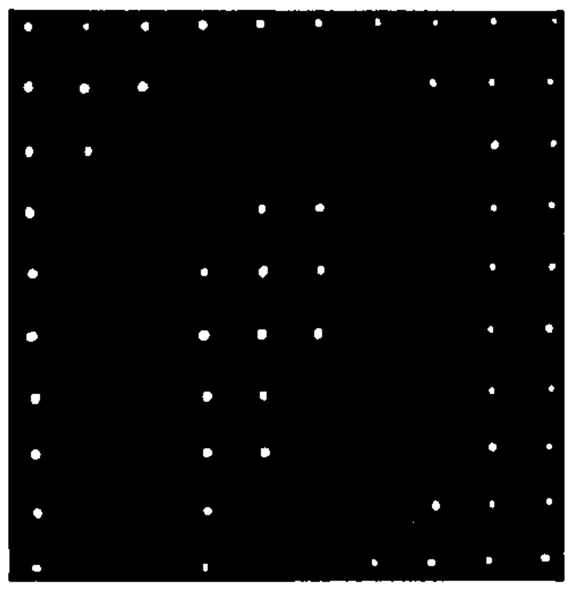
(b).



Output from Inverse Shuffle (image is split into two, horizontally sampled, sub-images of the loop).

Experimental :

(c).



(d).

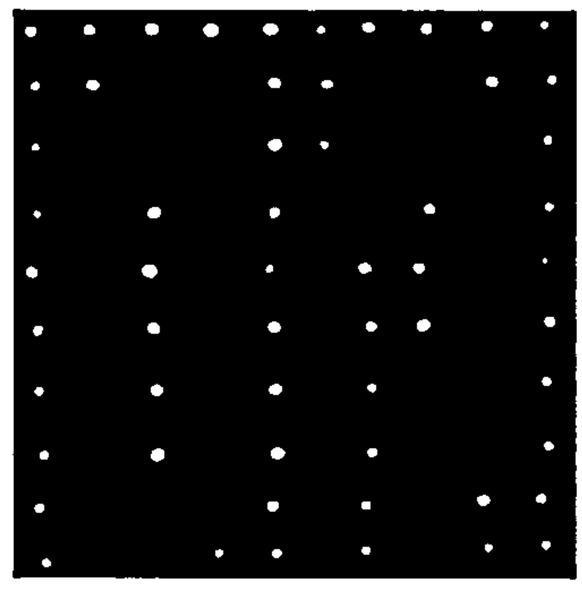


Figure 7.5: The performance of a 10×10 facet stacked 1D inverse perfect shuffle interconnect: (a). & (b). Idealised schematic of the input mask and output image; and (c). & (d). Photograph of the experimental results. SVHOE parameters are: $a = 500\mu$m, $f_h = 2$cm, $D = 5$cm, $\theta = 30°$, $\eta_{\text{singlet}} = 90\%$ and $\lambda_{\text{replay}} = \lambda_{\text{recording}} = 514.5$nm.

interconnected array. This is principally due to inaccuracies in the positioning of the holographic plate (at H) and the collimating lens (L_o) by the respective stepper motor systems (see figure 7.3); a problem that can easily be overcome by the use of better translation stages (see figure 7.9 in comparison).

7.1.4 Off-axis conclusions

Several off-axis SVIs based on a single pair of matched multi-faceted DCG holograms have been demonstrated for replay at the recording wavelength. However, the overall geometry of this approach makes it difficult to incorporate a wavelength-shift between recording and replay. In addition, the off-axis nature of the HOEs causes problems with alignment, polarisation and sensitivity to chromatic changes upon replay. Diffraction effects upon replay (see section 7.3) and the NA of the recording

optics will also limit the scalability of this technique.

7.2 On-axis SVIs

The implementation of an on-axis interconnect is illustrated in figure 7.1b, for a symmetric SVI. The basic unit of interconnection is a double element holographic array, shown in figure 7.2b. The doublet is made up of a holographic microlens array and a planar redirection array, the two volume holographic structures are aligned and cemented together to form the basic component, from which interconnects are built. The lens element collimates an array of (signal carrying) light beams of a fixed f/number into an array of plane waves at a fixed Bragg angle, (this off-axis deflection is needed to satisfy the spatial frequency requirements for an efficient volume hologram [LEI62]). The redirecting element receives the array of plane waves at a fixed Bragg angle and redirects them into free-space, in accordance with the spatial permutation defined by a specific interconnect. The cemented structure behaves as a single on-axis element. Point-to-point interconnection (between planes P_1 & P_2 of figure 7.1b) is achieved by using two doublet SVI components separated by the required permutation distance, D, so that the second SVHOE (H_3 & H_4), being a mirror image or inverse function SVI, performs the reverse operation of the first (H_1 & H_2).

7.2.1 Recording on-axis SVIs

Doublet on-axis SVI components, in which the collimation and redirection functions are independent, require the use of two recording systems to fabricate each component array separately. These geometries are illustrated in figure 7.6, they are both similar in layout, having the optic axes of object and reference arms at the same angles with respect to the holographic plate (same Bragg angle θ_B). Figure 7.6a is the basic recording geometry used to record high quality lenslet arrays, (with a focal length f_h), and is essentially the same as the off-axis recording scheme (see figure 7.3), without the $x-y$ stepper motor in the object arm. Figure 7.6b is based upon the recording systems used to produce planar grating HOEs with well defined

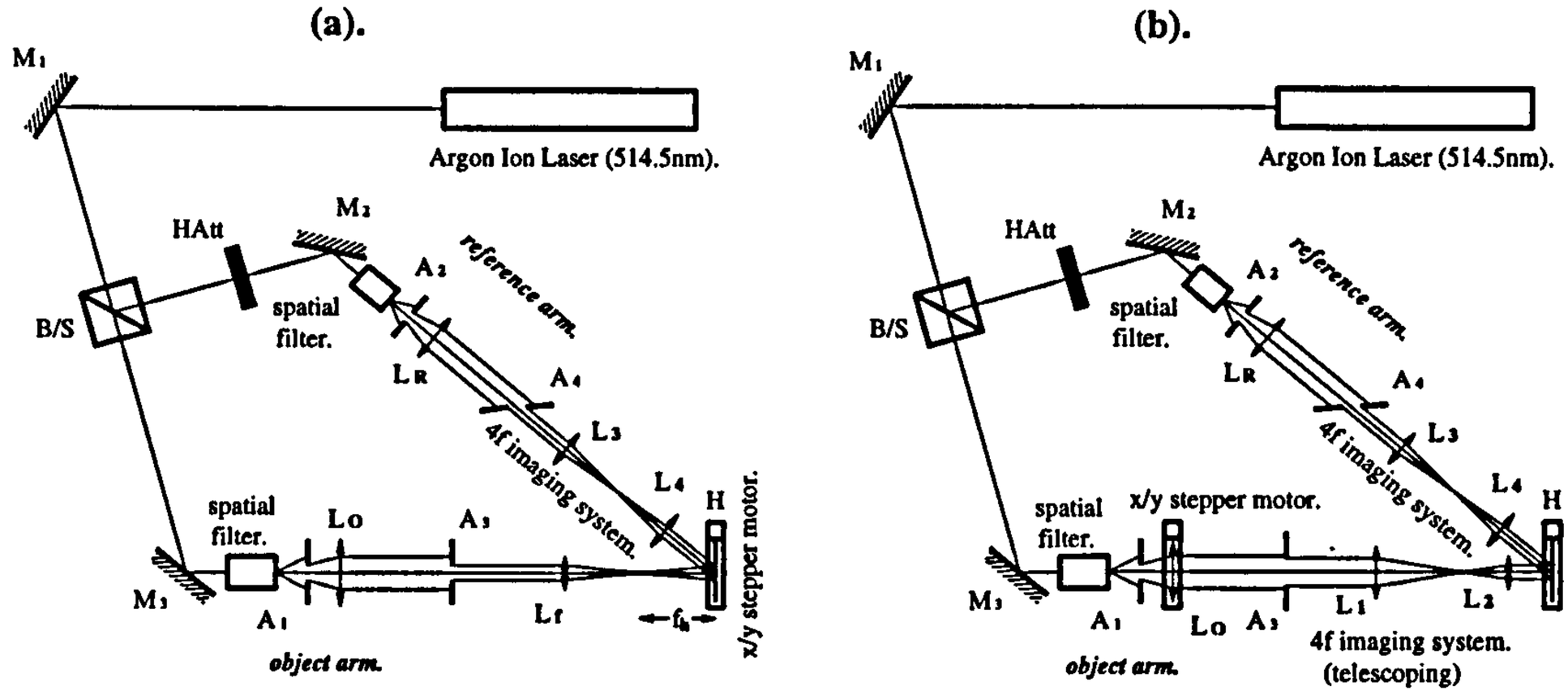


Figure 7.6: Recording systems used to fabricate on-axis (four element) doublet SVI components: (a). Lenslet/collimation element; and (b). Redirection/interconnection element, (the average Bragg angle, θ_B, between object and reference arms is the same between both set-ups :- 30°, $L_1 = 250$mm, $L_2 = 50mm$ and all optical components are labelled as per figure 7.3.)

edges, except that the object arm collimating lens, L_o, is mounted on an $x-y$ stepper motor. The on-axis recording systems contain similar optical components, making it easy to maintain the same Bragg angle between geometries by using the same set-up to record both doublet components and simply adding or removing lenses and repositioning certain elements in the object arm.

The technique used to record the planar grating array of redirection elements is the same as that described in section 7.1.1, for off-axis SVIs. The same equations relating lens translation and interconnection angle apply, except that since the object arm is perpendicular to the plate (i.e. the standard replay angle is zero), then $\theta = 0°$.

The same restrictions apply to this on-axis recording technique, that are outlined in section 7.1.1. The exception being that it is much simpler to incorporate a wavelength-shift into the design:

- The microlens collimating array is composed of lenslets with the same Bragg angle and is consequently relatively straightforward to design and record (see section 4.3); and
- The redirecting array is composed of planar gratings (with a common Bragg angle) that are even simpler to fabricate (see section 4.1).

Separating the collimating and redirecting functions eliminates many of the recording difficulties and potential geometrical aberrations associated with wavelength-shifting off-axis SVIs. Kobolla et al. [KOB94, KOB95] have fabricated efficient on-axis volume holographic space-variant interconnection components for replay wavelengths of 830nm and 1.5μm.

7.2.2 Replaying on-axis SVIs

On-axis SVIs are slightly easier to align than their off-axis counterparts, due in principle to the optical axis being common to both the external system and the interconnect. However, each array still has *six* degrees of freedom and consequently the same rigour must be applied to the opto-mechanical design and packaging of the interconnects.

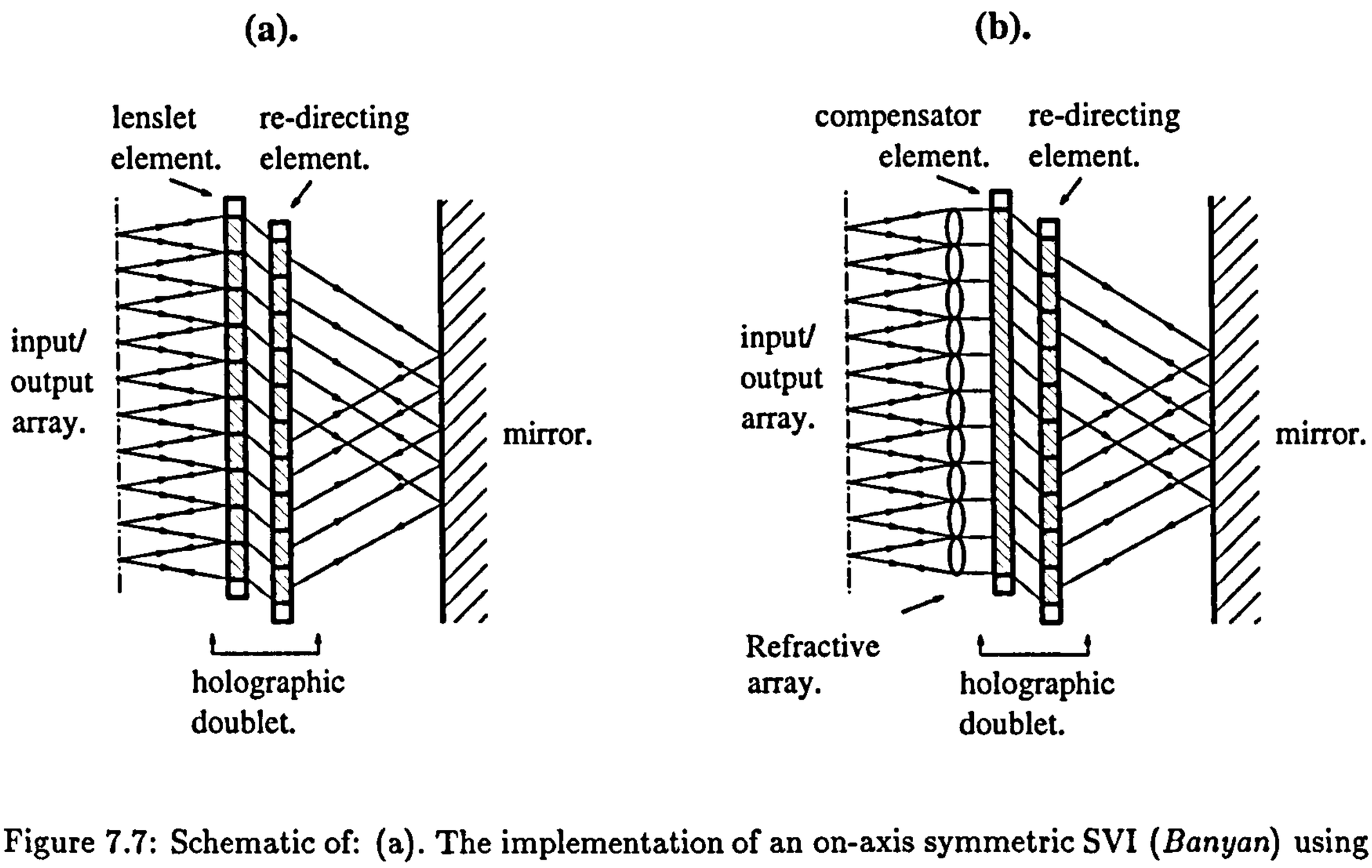


Figure 7.7: Schematic of: (a). The implementation of an on-axis symmetric SVI (*Banyan*) using a single HOE doublet and a mirror [RES91a]; and (b). The same SVI structure using refractive rather than diffractive microlenses [RES91b].

It is possible to simplify the alignment of symmetric on-axis SVIs, e.g. Banyan, because the pair of doublet SVHOEs are identical, (not inverse functions of one another). Consequently, the desired interconnect can be achieved using only one doublet SVHOE and a mirror placed at a distance $D/2$ (half-way), along the per-

mutation distance, as shown in figure 7.7a. The input and output arrays are spatially separated by a polarising beam splitter and quarter-wave plate combination (described in section 7.2.3). Eliminating the need for one of the SVI doublets halves the alignment difficulties, however, such a technique cannot be applied to asymmetric SVIs, e.g. the perfect shuffle.

Polarisation dependence, caused by different Fresnel transmission coefficients, is extremely low because all beams travel more or less normally to the doublet HOE surfaces, ($\Delta\theta_{max} \sim$ a few degrees).

Undiffracted light, (from the zero-orders of each SVI array), cannot propagate out of the system to the same extent as the off-axis geometry. However, Bragg mismatch between *four* separate HOEs make it unlikely that stray light will end up as noise in an output channel. Kobolla et al. [KOB95], using similar interconnects, have spaced each channel (with one pixel separation between them) and used arrays of stops in the inter-channel spacings in order to ensure that stray light does not leave the interconnect.

Close-cascaded HOE doublets are relatively insensitive to chromatic changes in the reconstruction wavefront [STO85, KOS89b] compared to singlet diffractive components. Consequently, on-axis SVIs are more robust to variations in replay wavelength [SCH92]. However, although the positional integrity of the output array can be maintained and an acceptable interconnect efficiency (with minimal crosstalk) is achievable using pairs of doublet SVIs, the variation in focal length (of diffractive elements) due to dispersion, is not so easily compensated. A change in wavelength $\Delta\lambda \sim \pm$20nm, will bring about a corresponding focal length variation $\Delta f \sim \pm$2.3%, for diffractive elements at 850nm. It is possible to obviate this problem by the use of refractive collimating micro-lens arrays, as illustrated in figure 7.7b. This approach still utilises a close cascaded HOE doublet (with the inherent chromatic tolerance benefits), however, the first element is a planar grating compensator that couples the channels into the redirecting array at the correct Bragg angle. Sauer et al. [SAU94] have successfully demonstrated on-axis refractive-diffractive micro-optic SVI elements using surface relief rather than volume HOEs. A similar wavelength variation to that above, $\Delta\lambda \sim \pm$20nm, results

in $\Delta f \sim \pm 0.005\%$ for their refractive microlenses. Since McCormick et al. [MCC91] have demonstrated that the distance P_1-to-H_1 is a critical factor in SVI performance, the refractive-diffractive combination would significantly improve the replay tolerances of the interconnect in this respect.

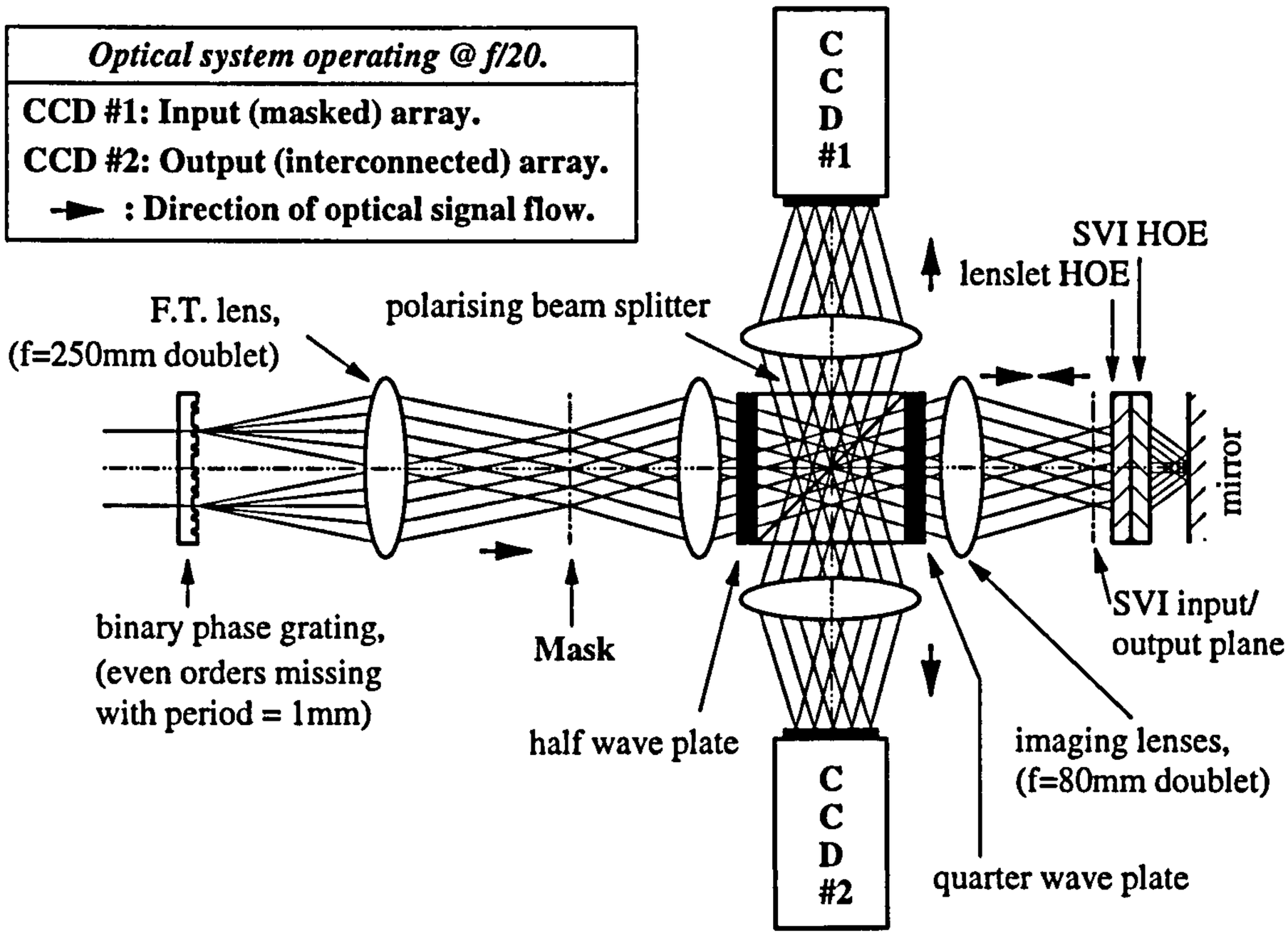


Figure 7.8: Optical system used to test the replay of an on-axis, symmetric SVI doublet used in a reflective orientation. CCD#1 records an image of the input mask and CCD#2 an image of the interconnected spot pattern, (ray paths are shown in 1D for a five channel interconnect).

7.2.3 Experimental results

As with the off-axis interconnects, a range of on-axis SVIs were fabricated to demonstrate their feasibility with regard to shuffling networks, shuffle-equivalent networks and arbitrary one-to-one interconnects. Full 2D permutations were generated between input and output facets in the $x-y$ plane. All of the holograms were recorded and replayed at 514.5nm using DCG derived from Kodak 649F plates and standard pre-processing and development procedures, (see section 2.7).

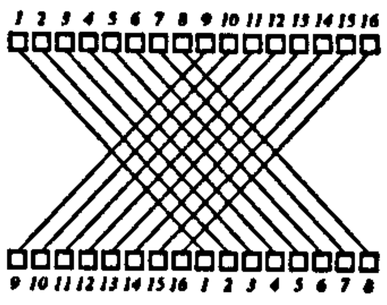
Figure 7.8 shows the optical system used to demonstrate the replay of an on-axis symmetric SVI doublet in the reflective orientation (see section 7.2.2 and figure 7.7a).

A binary phase grating (BPG) is illuminated by a plane polarised, collimated beam (at 514.5nm) and generates a 16 × 16 array of spots in the mask plane. The mask plane is $4f$ imaged through the polarising beam splitter (PBS) onto the SVI input/output plane. The half-wave plate is rotated so that some of the light from the mask plane becomes $s-$polarised and is then $4f$ imaged onto CCD#1 in order to monitor the masked input to the SVI. The on-axis SVI doublet and mirror combination permute the beamlets according to the space-variant operation (in this case a 2D first-stage Banyan) and the interconnected spot array is formed at the SVI input/output plane. This plane is then $4f$ imaged onto CCD#2 to monitor the interconnected output of the SVI. Input to and output from the interconnect are split by the use of the quarter-wave ($\lambda/4$) plate: the $p-$polarised light from the mask plane is imaged straight through the PBS, passes through a $\lambda/4$ -plate (with its fast axis at 45°with respect to the PBS plane of incidence) presenting circularly polarised light to the volume HOEs, after interconnection the signal array passes once more through the $\lambda/4$ -plate becoming $s-$polarised as it enters the PBS and is thus imaged onto CCD#2. This is essentially the optical system required to incorporate these SVIs into demonstration circuits (see section 7.4).

Figure 7.9 illustrates the performance of a 16 × 16 facet 2D first-stage Banyan interconnect when used in the optical system shown in figure 7.8. Figures 7.9a & b show the idealised input mask and interconnected output to and from the SVI, respectively. The 1D Banyan, when mapped into 2D, performs the same operation on a 2D image as a stacked 1D interconnection horizontally followed by one vertically. The net effect is to swap diagonally opposed quadrants. Figures 7.9c & d show the resulting experimental input and output of the interconnect, observed on CCD#1 and CCD#2 in figure 7.8. Note that the central bright spot in the input plane is due to the BPG being etched to the incorrect depth for efficient replay at 514.5nm, however, it is not present in the output array.

In this particular demonstration the facet size, a, was 250μm, the interconnection distance, D, was 5cm, the Bragg angle, θ_B, between doublet components was 30° and the replay focal length, f_h, was 5mm at 514.5nm. The optical system is consequently operating at $f/20$ - a figure that was stipulated by the availability of the 16 ×

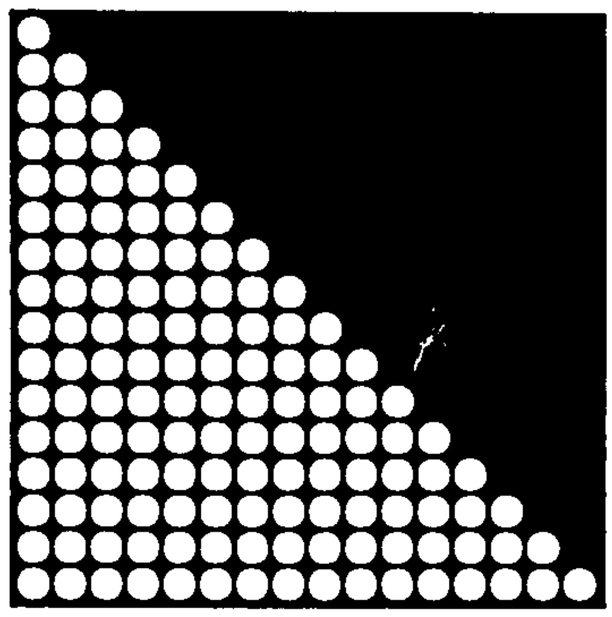
Schematic :



1D Bantan (1st stage) schematic for 16 facets.

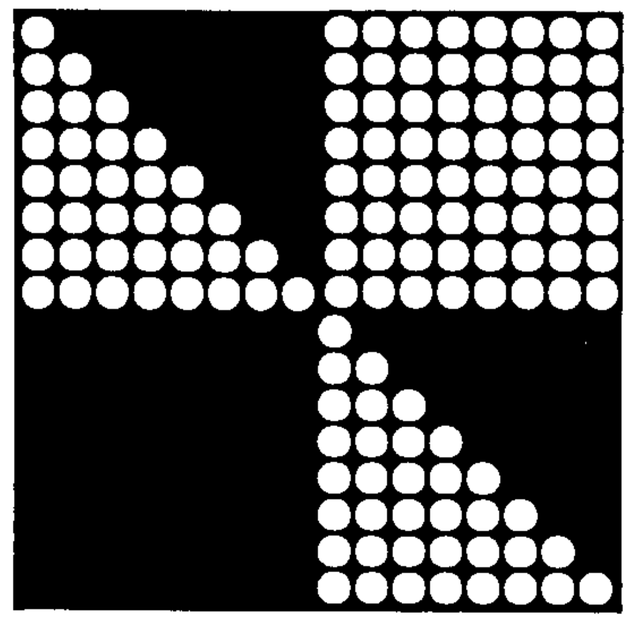
Experimental :

(a).



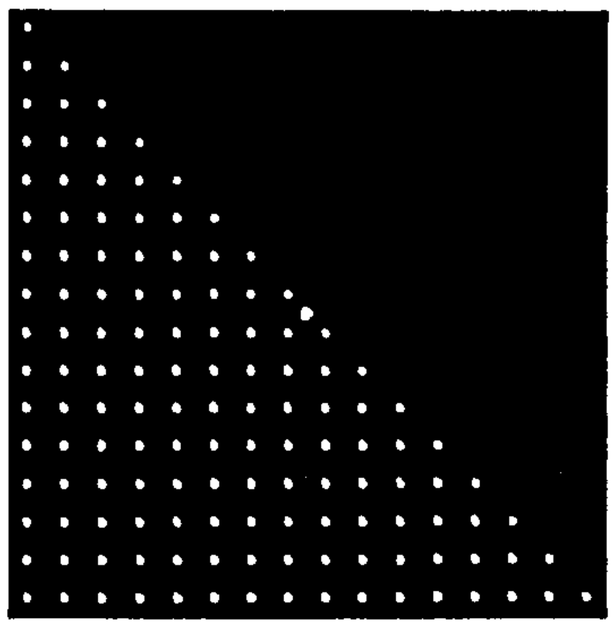
Masked input to Banyan (mask split diagonally accross the input field).

(b).



Output from Banyan (all four quadrants swapped in 2D at once).

(c).



(d).

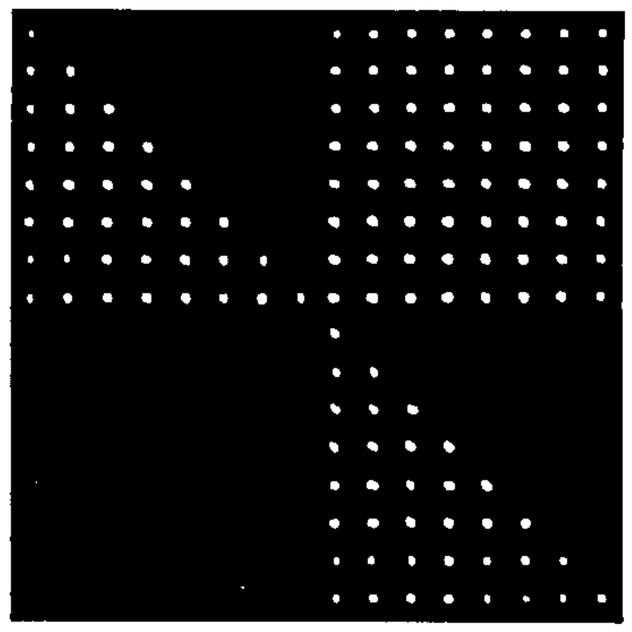


Figure 7.9: The performance of a 16×16 facet 2D 1st-stage Banyan interconnect, (simultaneously implements the 1D pattern in two mutually orthogonal directions): (a). & (b). Idealised schematic of the input mask and output image; and (c). & (d). CCD images of the experimental results, (corresponding to the spot patterns observed on CCD#1 and CCD#2 in figure 7.8). SVHOE parameters are: $a = 250\mu$m, $f_h = 5$mm, $D = 5$cm, $\eta_{\text{doublet}} = 87.6\%$ and $\lambda_{\text{replay}} = \lambda_{\text{recording}} = 514.5$nm.

16 BPG - faster systems are possible, simply by changing f_h of the collimating array. The efficiency of the SVI doublet was measured to be 87.6%, giving the overall interconnect a diffraction efficiency of 76.7%, (neglecting Fresnel losses).

7.2.4 On-axis conclusions

A range of on-axis SVIs with full 2D mapping between channels have been demonstrated, based on a matched pair of doublet DCG holograms replayed at the recording wavelength. This approach is much easier to wavelength-shift than off-axis SVIs and being a close-cascade, on-axis geometries are more tolerant to alignment, chromatic, polarisation and aberrative effects. Chromatic tolerances can be increased still further by the use of refractive microlens arrays as collimating/concentrating

elements. The scalability of the technique is, however, limited (in the same way as off-axis SVIs) by diffraction effects upon replay (see section 7.3) and the NA of the optics used to record the interconnects.

7.3 Design considerations

There are several key factors that must be taken into account when designing SVIs, they have been analysed in detail by a variety of researchers; an overview is presented here.

7.3.1 Diffractive effects

SVIs for use in practical circuits need to have facet sizes of between 100 and 300μm. Clearly, where such small beam diameters are concerned the effects of diffraction must be considered. If this is not done at the design stage then, on replay, signal beams will diffract over the permutation distance and end up overfilling the facets of the second array of SVI elements. This leads to power loss (reduced η) and an increase in cross-talk between pixels (decreased SNR), both of which can fundamentally limit the performance of the interconnect. Figure 7.10 illustrates a diffraction compensated interconnection for incident beams with the same Gaussian intensity profile. This figure shows the off-axis geometry, but is equally applicable to on-axis interconnects. The only difference being that off-axis SVIs produce an astigmatic wavefront [HAL89] when diffracted by SVHOE#1, this can be avoided with the on-axis geometry. The figure shows the ideal symmetric-beam case where the beam waist, 2ω, at the input plane is equal to that at the output plane (for collimating elements of the same focal length, f). This implies that the beam diameter at SVHOE#1, $2\omega_H$, is equal to that at SVHOE#2, (note that $2\omega < a$, this is discussed in section 7.3.2). In this ideal situation, the intermediate beam waist, $2\omega'$, is located exactly half-way along the permutation distance, D. Gaussian beam propagation theory [HAL89] shows that the criterion required to comply with this symmetric-beam condition is $D = 2f$. Clearly the effects of diffraction, limit the interconnection distance, D, for maximum power throughput. Limiting the al-

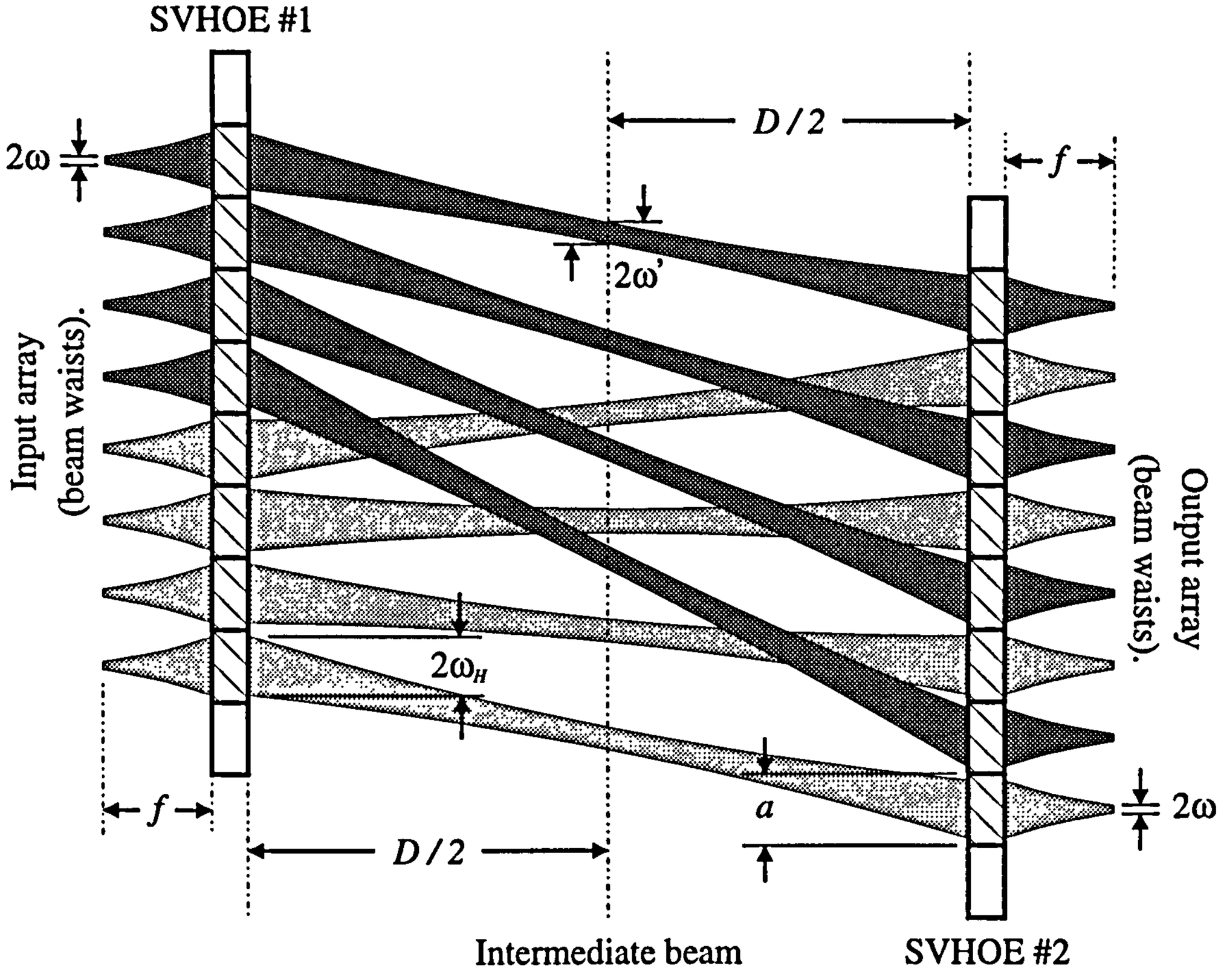


Figure 7.10: Gaussian beam propagation, (irradiance e^{-2} waist radius, ω, at the input & output planes), through an off-axis volume SVI, (8 facet perfect shuffle), showing the ideal location of beam-waists (e^{-2} radius, ω') and the underfilling requirement for HOE apertures to maximise throughput and minimise crosstalk, (ideally $3\,\omega_H \leq a$). Similar conditions apply to on-axis SVIs.

lowable interconnect distance, D_{max}, requires larger interconnection angles $\Delta\theta_{max}$ for the SVI elements; with the recording scheme presented here this is undesirable. Zaleta et al. [ZAL95] analyse the trade-off in fractional power falling on SVHOE#2 as a function of permutation distance, D, for various HOE aperture sizes, a.

It is possible to overcome this diffraction limit, imposed upon D_{max}, in one of two ways:

1. **Defocusing into the interconnect.** McCormick et al. [MCC91] have analysed the Gaussian beam propagation problem for on-axis refractive lenslet arrays and have looked at two distinct situations: the first being the ideal collimation case, outlined above, in which minimum beam divergence is

paramount; the second is a defocused condition in which the lens to intermediate waist distance is maximised. The defocused situation allows the distance between the input plane and SVHOE#1 to increase by a factor upto the Rayleigh range, z_R, of the source/micro-lens. This increases the distance over which the intermediate beam waist, $2\omega'$, is relayed and allows D_{max} (equal to $2f + f^2/z_R$, [MCC91]) to be increased. Care must be taken not to overfill the lens aperture.

2. **Build focusing power into the permuting beams.** There are two possibilities to consider, both of which aim to increase D_{max} whilst maximising power throughput and minimising cross-talk over the simple collimated beam geometry:

   - Incorporate focusing power into the redirecting element of SVHOE#1 such that at SVHOE#2 the diffraction limited spot of the permuted beam is roughly the same size as the aperture of SVHOE#2; or
   - Incorporate focusing power into both SVHOEs so that the permuted beams focus down to a beam waist at $D/2$.

   The first design has been demonstrated by Urquart et al. [URQ94] for off-axis Fresnel zone plate (FZP) SVIs. This solution to the limit on D_{max} makes the recording of both the off- and on-axis elements, discussed in this chapter, much more complicated where a wavelength-shift (see chapter 4) is concerned.

Considering the two demonstrations shown in figures 7.5 & 7.9, a defocusing upto the Rayleigh range in both cases implies a change in the collimating distance of +5.2%, this allows for a D_{max} of upto 422mm and 105mm respectively, at full defocusing. Clearly this solution is preferable to building focusing power into the permutation elements.

7.3.2 Aperture filling

Consider the SVI geometry shown in figure 7.10. Since the intensity profile of a Gaussian beam extends to infinity it is clear that a fraction of the incident power

could spill over into adjacent pixels and the beam will be clipped at the collimating element. In order to minimise the effects of beam-clipping an evaluation of the way that power coupled into the HOE varies with the ratio of the Gaussian (e^{-2} intensity) beam radius, ω_H, (at the hologram) to the collimating aperture size, a, (for square facets) is required.

The total power incident on a facet, P_F, is given by

$$P_F = 4\int_0^{a/2}\int_0^{a/2} I(x,y)dxdy, \tag{7.4}$$

where $I(x,y) = I_o \exp\left[-\frac{2(x^2+y^2)}{\omega_H^2}\right]$. This reduces to

$$P_F = \frac{\pi I_o\, \omega_H^2}{2}\left[\mathrm{erf}\left(\frac{a}{\sqrt{2}\,\omega_H}\right)\right]^2 \tag{7.5}$$

where $\mathrm{erf}(z)$ is a standard error function defined as

$$\mathrm{erf}(z) = \frac{2}{\sqrt{\pi}}\int_0^z \exp[-t^2]dt \tag{7.6}$$

Since the total power in the incident Gaussian beam, P_T, is given by $P_T = \frac{I_o \pi \omega_H^2}{2}$, then the fraction, F, of the incident light that ends up in the correct interconnection channel can be expressed as

$$F = \frac{P_F}{P_T} = \left[\mathrm{erf}\left(\frac{a}{\sqrt{2}\,\omega_H}\right)\right]^2 \tag{7.7}$$

Figure 7.11 illustrates the dependence of this fill factor, F, on the ratio a/ω_H.

Light that overfills the facet will not necessarily diffract along the interconnect paths of the nearest neighbour elements, because the Bragg condition will not be correctly satisfied. This minimises the possibility of cross-talk as the light will remain in the zero-order and not be interconnected, as noise, to a spot in the output plane.

Clipping does, however, have two major effects:

- **Reduces the overall efficiency of the interconnect.** It can be seen from figure 7.11 that for $a \simeq 2\omega_H$, less than 90% of the incident light is coupled into

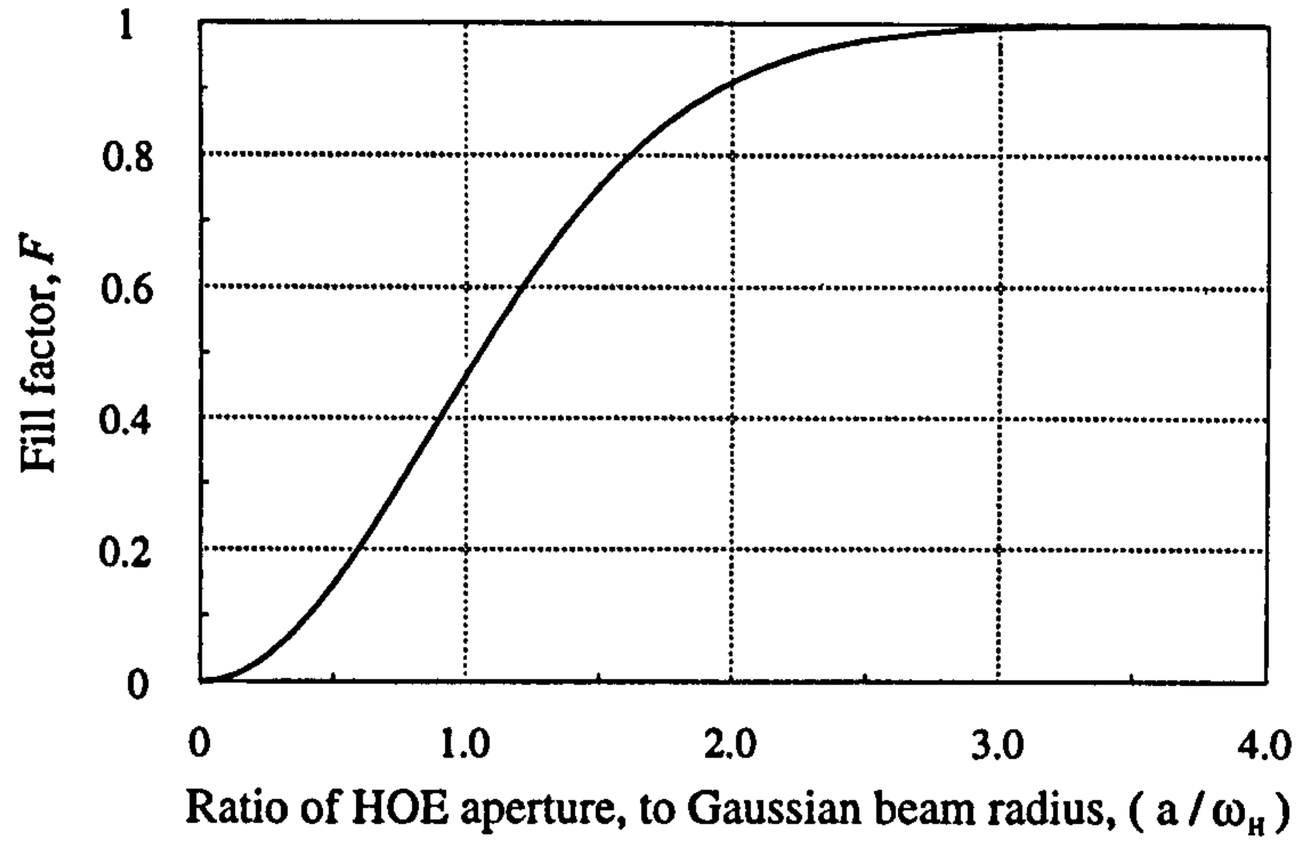


Figure 7.11: Fill factor, F, variation versus facet size, a, to Gaussian beam radius, ω_H, ratio. This represents the fraction of the incident light that ends up in the correct interconnection channel for SVIs of differing aperture.

the interconnect. However, for $3\omega_H \leq a$, over 99% of the power ends up in the correct facet. Consequently, the condition $3\omega_H = a$ is a useful criterion to use when designing such SVI systems.

- **Changes the beam propagation characteristics.** It has been shown by McCormick et al. [MCC91] (for circular apertures) that the change in intermediate beam waist size and location can be quite significant, e.g. clipping caused by $2\omega_H = a$ alters these parameters by up to 5%. However, in the region $3\omega_H \leq a$, these effects are insignificant and so the design criterion of $3\omega_H = a$ is reinforced.

Thus to minimise the effects of clipping of Gaussian beams, on both power throughput and propagation characteristics, the condition $3\omega_H = a$ should be met.

7.3.3 Packing density and scaling

The effects of diffraction, discussed in section 7.3.1, lead to an additional restriction pertaining to the achievable packing density of hologram facets in a SVI array and consequently affect the overall size and scalability of the interconnect. This problem has been analysed by Feldman & Guest [FEL89] for CGHs and by Urquart et al. [URQ94] for off-axis (FZP) SVIs.

A discussion is presented here for the three volume hologram/refractive microlens configurations discussed in section 7.3.1:

- Exact collimation of permuting beams,
- Defocused collimation of permuting beams; and
- Focusing power built into permuting elements.

To estimate packing densities for off- and on-axis geometries, conforming to the Gaussian model of figure 7.10, a few additional simplifications are necessary: HOEs are assumed to be thin compared with the interconnection length, facets are square and the maximum interconnection angle is less than 45°, i.e. off-axis $\theta+\Delta\theta \leq 45°$and on-axis $\Delta\theta \leq 45°$. For a maximum interconnection distance, D_{max}, this means that

$$aN \leq D_{max}, \tag{7.8}$$

where a is the facet size and N is the linear number of facets in a $N \times N$ SVI array, (note that the tilting mirror recording technique [ROB88, HAU89] is more suited to these extremes of interconnection angle than the lens translation presented here).

Exact collimation

In the case of exact collimation for each channel i.e. the distance P_1-to-$H_1 = f$, then $D_{max} = 2f$. This limits the linear size of the SVI array to $2f$, consequently to achieve a large number of channels, N^2, high f/number optics ($\sim f/30$) are required, e.g. a 142×142 array of 350μm aperture HOEs requires $D_{max} = 5$cm, which is $f/71$. Such slow optical systems are beneficial in terms of alignment and aberrations, however, high quality imaging optics are required to change between the f/number of the external system ($\leq f/5$) and that of the SVI.

Defocused collimation

Defocusing between P_1 and H_1 so that the maximum distance is $f + z_R$ implies that

$$2f \leq D_{max} \leq 2f + \frac{f^2}{z_R}, \tag{7.9}$$

where the lower limit equates to the exact collimation condition discussed above. For a Gaussian beam of e^{-2} radius ω_H at H_1 and taking into account the fill factor requirement (see section 7.3.2) of $a = 3\omega_H$, then the Rayleigh range z_R, can be expressed as

$$z_R = \frac{9f^2\lambda}{\pi a^2} \tag{7.10}$$

Substituting this into equation 7.9 gives the minimum allowable facet size to avoid diffractive cross-talk:

$$a_{min}^2 \geq \frac{9\lambda}{\pi}(D_{max} - 2f) \tag{7.11}$$

This can be used together with equation 7.8 to evaluate minimum facet size and maximum number of interconnection channels, N_{max}^2, for a fixed interconnection distance of D_{max} and collimating elements of focal length f.

Table 7.1a lists pixel sizes and the number of interconnects for SVI arrays of linear dimension D_{max} at a replay wavelength of 850nm. Those elements left blank correspond either to exact collimation or break the assumption expressed in equation 7.8. The interconnect density is comparable to exact collimation, however, faster microlenses make the SVI more compatible with the external optical circuit, e.g. D_{max} = 5cm for a facet size of 342μm and collimating optics of f = 1mm supports a 146 × 146 array at $f/2.9$.

D_{max} (mm)	a_{min} (μm)			N_{max}^2		
	f 250μm	f 1mm	f 4mm	f 250μm	f 1mm	f 4mm
1	35	-	-	821	-	-
2	60	-	-	1095	-	-
5	105	85	-	2281	3422	-
10	152	140	70	4322	5133	20533
50	347	342	320	20740	21388	24444

(a). Defocused collimation of Gaussian beams

D_{max} (mm)	a_{min} (μm)	N_{max}^2
1	62	261
2	87	522
5	138	1307
10	196	2614
50	437	13071

(b). Focused permuting beams

Table 7.1: Scaling of SVIs with varying interconnection distances D_{max}, for : (a). the defocused collimation condition of McCormick et al. [MCC91], with different collimating element focal lengths, f; and (b). focusing power in the permuted beams, after Schwider et al. [SCH92]. The results of Schwider et al. are modified for the fill factor condition $a = 3\,\omega_H$.

Built-in focusing power

The diffraction density limits of building focusing power into the beam permuting side of the redirection element have been analysed in similar ways by Derstine & Guha [DER89] and Schwider et al. [SCH92]. In this case, the intensity distribution at the second SVHOE plane, due to the focusing properties of the first SVHOE, can be calculated using Fraunhofer approximations. The intensity distribution has a $sinc^2$ character, so assuming that the distance between the first positive and negative minima of the diffraction spot, at SVHOE#2, is the same as the input beam diameter at SVHOE#1 ($2\omega_H$) and taking into account the aperture filling requirement (see section 7.3.2) of $a = 3\omega_H$, then:

$$a_{min}^2 = \frac{9\lambda D_{max}}{2} \tag{7.12}$$

Again this can be used with equation 7.8 to evaluate the maximum number of channels supported, N_{max}^2. Table 7.1b lists pixel size and number of channels for the same linear dimensions, D_{max}, and wavelength as the defocused collimation case, (the *original* results of Schwider et al. [SCH92] do not include the $a = 3\omega_H$ requirement). It can be seen from the two tables that the focusing power solution reduces the number of supported channels by about half.

Finite thickness considerations

When the finite thickness of a DCG HOE is taken into account it affects packing density in two ways:

1. Incomplete overlap between the object and reference beams through the depth of the medium, d, during recording reduces the effective facet size. For example, recording angles of 0° and 30° with $d = 15\mu$m reduces the facet size by $\sim 5\mu$m; and

2. As the beam waist, $2\omega_H$, approaches the hologram thickness the diffracted beam becomes distorted due to the incomplete conversion of the zero-order beam into the first diffracted order [CHU77b]. This places a lower limit on

facet size and an upper limit on the f/number of the collimating optics (over and above those already discussed), since $\omega_H > d$ which implies $a > 3d$. However, for most SVI systems the facet size is between 100μm and 300μm so this is not a major concern.

Density conclusions

Although the packing density of these SVI systems is not comparable to that of space-invariant free-space optics, (*cf.* the resolving power of a lens being $\sim 10^6$ channels), tens of thousands of complex interconnections can still be squeezed into an area the size of an optoelectronic chip.

7.4 SVIs in optical circuits

Schwider et al. [SCH92] discuss various ways of incorporating on- and off-axis DCG and CGH SVI geometries into photonic switching networks using in-line optics. A brief explanation of how on-axis SVIs can easily be incorporated into, SEED and NLIF based, photonic switching fabrics follows.

7.4.1 SEED devices

Figure 7.12a shows an optical module that can be used with SEED based switching devices to build multi-stage interconnection networks. The configuration shown is for on-axis symmetric SVIs operating in a reflective orientation (as per figure 7.7a) however, it can be used with asymmetric interconnection patterns (see figure 7.14). In figure 7.12a, the optical arrangement around the left-most beam splitter deals with the interconnect whilst that around the right-most delivers the interconnected channels and power/control signals to the SEED. The $\lambda/4$-plate and PBS combinations operate in the same way as described in sections 7.2.2 & 7.4.3 and by Hinton et al. [HIN94]. The patterned mirror reflector (PMR) arrays are required to space-multiplex all of the input signals onto the SEED windows [HIN94] while enabling the output signal to be relayed to the next module, (SEEDs can be optically

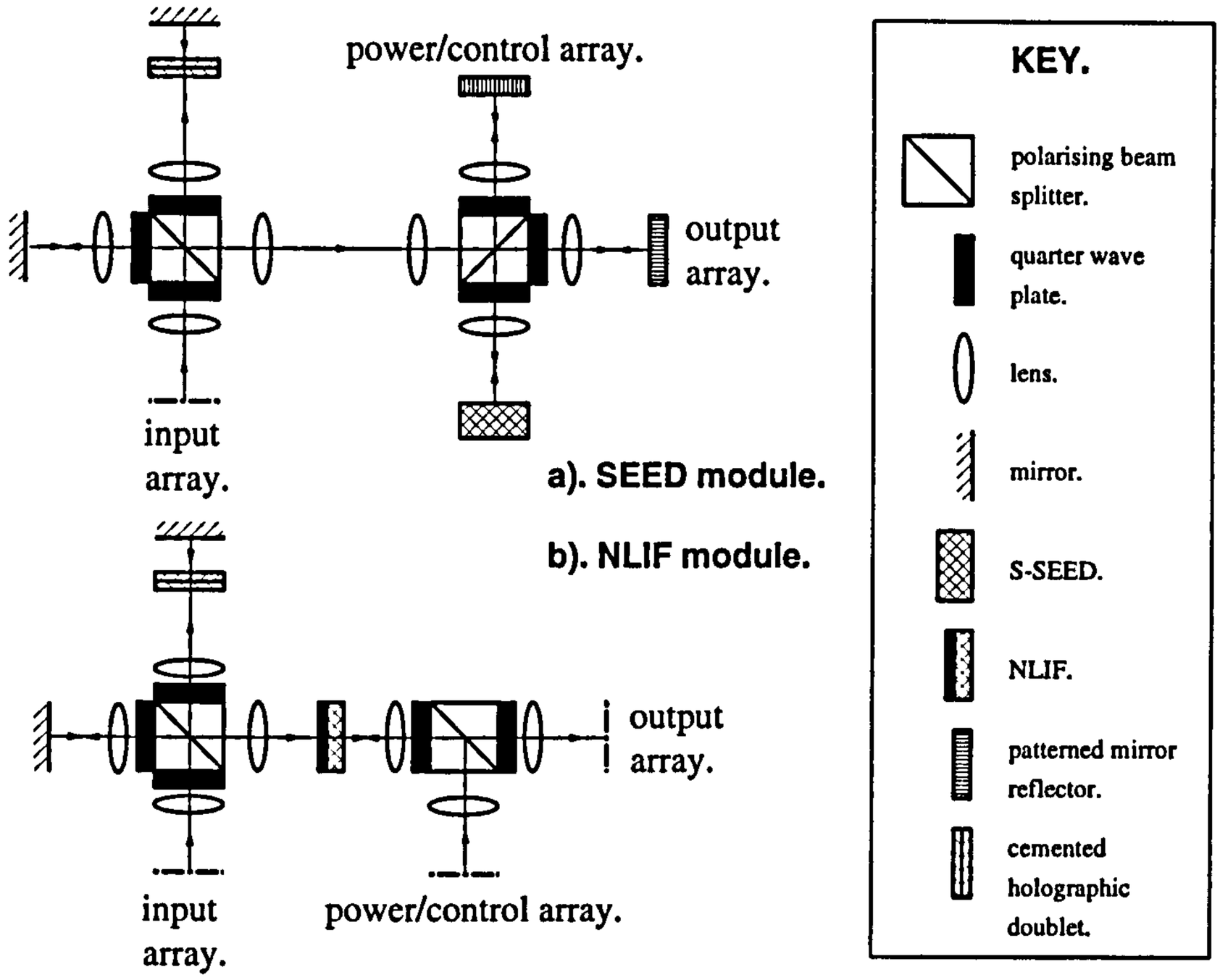


Figure 7.12: Modular optics for the use of on-axis SVIs in optical circuits compatible with: (a). SEED and (b). Fabry-Perot (e.g. NLIF) switching devices.

addressed from one side only).

7.4.2 Fabry-Perot devices

Unlike SEEDs, Fabry-Perot based switching devices such as the NLIF [BUL89] can be optically addressed from both sides. This makes it much simpler to deliver all of the control, power and signal beams, (PMR arrays are not required). Figure 7.12b shows the equivalent optical module to implement multi-stage networks using bistable-etalons with absorbed transmission (BEAT) devices.

7.4.3 Simplifying complex interconnects

Of all the regular multi-stage SVIs, the perfect shuffle is the most general in its interconnection and the only one that is stage-invariant. When used in shuffling networks it is composed of two sub-shuffling stages, shown in figure 7.13, which for the sake of clarity will be called the perfect shuffle and the stacked deck. Figure 7.13 illustrates the relationship between the perfect shuffle and the stacked deck; the

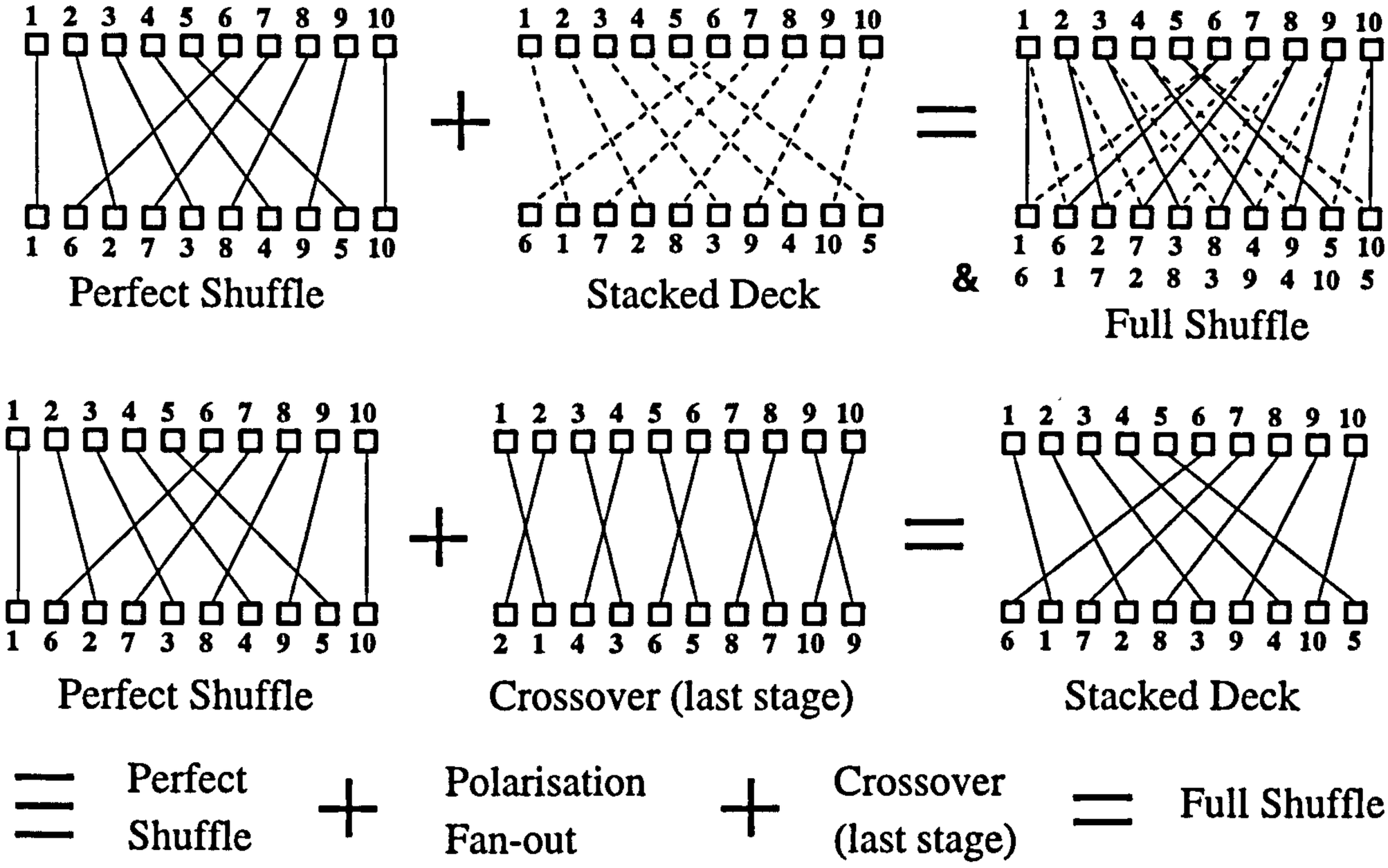


Figure 7.13: Schematic illustration of how a full shuffle interconnect can be implemented using an asymmetric (*perfect shuffle*) and an symmetric (*crossover*) interconnect, rather than the more complicated twin asymmetric system, (*perfect shuffle* and a *stacked deck*). In circuits where switching devices can perform exchange-bypass operations (e.g. SEED based 2-modules [HIN94]), the crossover can be disposed with.

latter being equivalent to the former followed by an exchange between neighbouring facets.

SEED based switching devices can perform exchange-bypass operations, which eliminates the need for a stacked deck SVI, in order to implement the full shuffle. However, where devices that do not perform such operations are used, e.g. NLIFs, the need for two complex asymmetric SVIs can be overcome. Figure 7.14 shows how a perfect shuffle can be combined with a crossover interconnect (used in a reflective orientation - see section 7.2.2) to achieve a full shuffling stage. The PBS and $\lambda/4$ - plate combinations are used to split the perfectly shuffled array $^{50}/_{50}$ between the crossover and the straight-through (retroreflected) interconnection and to recombine them in the output plane (as described in section 7.2.3).

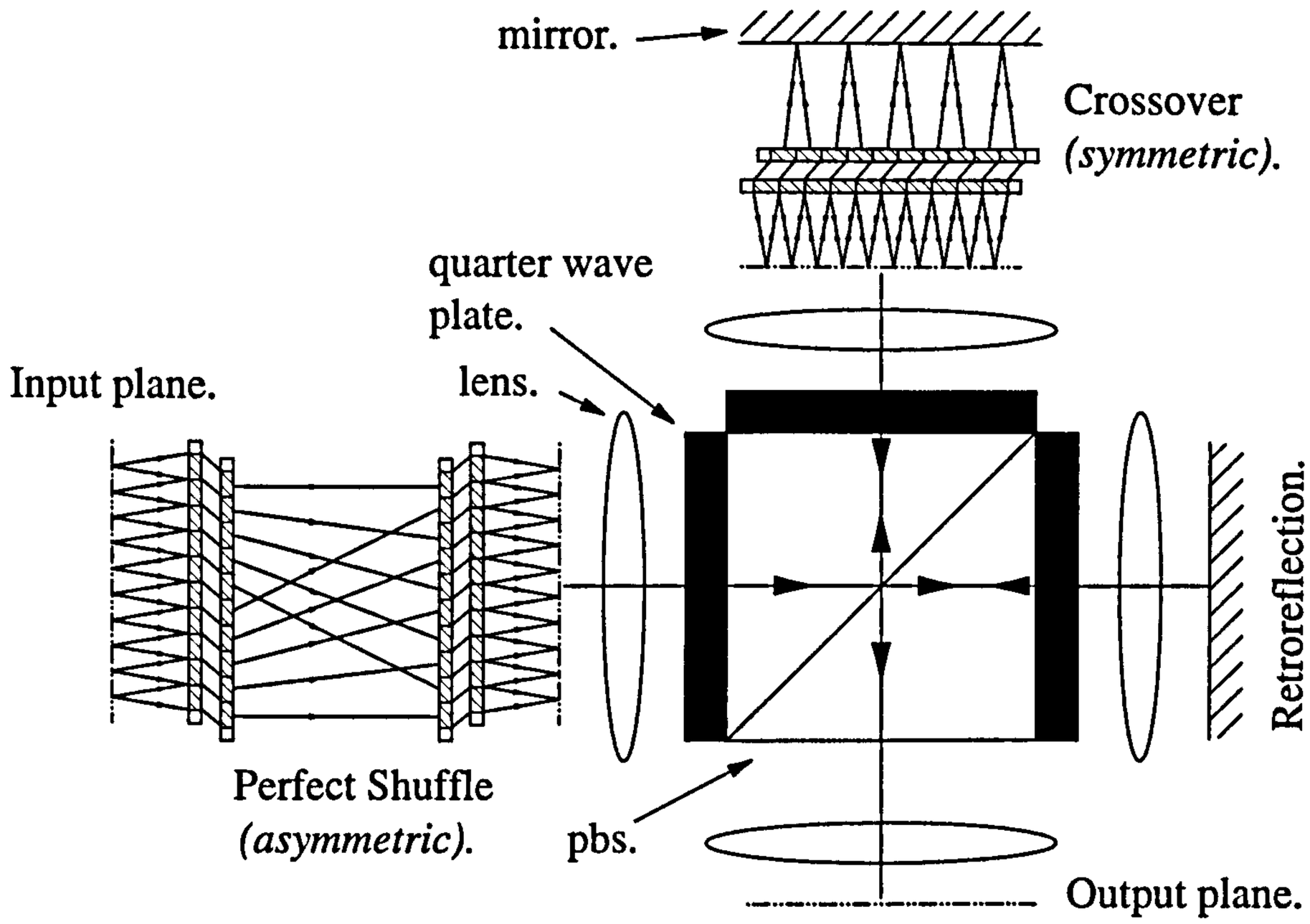


Figure 7.14: Optical realisation of a full shuffle using on-axis perfect shuffle and crossover SVIs, (path lengths through the crossover and retroreflected interconnects are made equal to eliminate clock-skew) [RES91d].

7.5 Summary

Both off- and on-axis SVIs have been demonstrated for replay at the recording wavelength. The design, recording and replay considerations discussed, coupled with the experience of aligning the two geometries, lead to the conclusion that the on-axis system is preferable.

Although this volume HOE fabrication technique has the capacity to provide functional SVIs with a high efficiency and similar interconnects have been successfully incorporated into demonstration optical circuits [GLU93], the limitations of the approach have still to be defined; particularly with regard to the accuracy of fabrication, packaging and practical demonstrations of scalability.

Highly efficient SVIs are feasible, can be designed for use at current communications wavelengths [KOB94, KOB95] and consequently have the potential to be more useful than the restrictive connectivity of SIIs.

Chapter 8

Conclusions

The theoretical and experimental work carried out in this thesis has shown that volume holographic optical interconnects are a viable technology in the field of optical signal processing.

A reliable fabrication method has been described for the manufacture of highly efficient (> 90%) transmission and reflection HOEs in DCG. Along with this, methods of protecting and fine tuning holograms have been explained that can produce HOEs that meet the positional, efficiency and uniformity tolerances of experimental optoelectronic information processing systems.

A new and improved rigorous couple-wave theory has been presented that is capable of modelling absorption and/or phase modulated holograms, including grating harmonics. This model has been used in the study of both transmission and reflection interconnects. Comparisons with a multi-wave model [RED89a] have suggested that grating non-uniformities in DCG emulsions may not be as large as previously thought.

It has been shown that the recording restrictions imposed by the limited spectral sensitivity of DCG can be overcome using a new simplified wavelength shifting technique. Equations used to calculate wavelength-shifted recording geometries for both planar gratings and those with focusing power have been presented. An analysis of the technique, using ray-tracing models [RED89a, RED89b, ORA94], showed that the performance of such wavelength-shifted lenses was the best that could be

achieved using spherical recording wavefronts. The method produced microlenses (< 1mm aperture) that operated close to the diffraction limit for lenslets of $f/3$ or greater, shifted from the green recording wavelength to the near IR replay. However, where faster optics or larger lenses are required aspheric recording wavefronts are necessary, e.g. [CHE87, ASS88].

The rigorous coupled-wave theory has been further extended to model the general case of multiply superimposed planar gratings in reflection and transmission. This new theory was used to analyse the behaviour of optically recorded fan-out gratings. Quantitative agreement was obtained with the experimental and theoretical work of Redmond [RED89a, chapter 4] for transmission gratings. Similar agreement was reached with the work of Kostuk et al. [KOS86] for reflection fan-out to two. A more detailed theoretical analysis of higher level fan-out using reflection elements showed that (unlike transmission gratings) multiple-grating interactions need not severely limit efficiency and replay fidelity. Indeed, it has been shown that with reflection gratings it is theoretically possible to achieve $\sim$ 78% useful diffraction efficiency with $\sim$ 1% uniformity variation, without resorting to the grating phase control techniques required for transmission elements [HER92]. It is believed that this is the most generally applicable model to date for the analysis of optically recorded volume fan-out holograms.

A number of experimental spatially variant and invariant interconnects have been fabricated, for operation in the visible and near IR, using the techniques described in the thesis. These ranged from bus-type connections, through shuffle networks to crossbar systems. All of these holographic interconnects have been demonstrated as a feasible solution to the interconnection problem that they were designed to address. The only drawback, that is common to all, is the limited fabricational accuracy of optically recorded holograms. Greater control over the positioning of optical components during the recording process would overcome this problem.

At the time of submission DCG is no longer the most dominant volume holographic material, having greater competition from DuPont and Polaroid photopolymers. Indeed, volume holographic gratings have, for the time being, been largely displaced by surface-relief gratings (e.g. [MIL93]) as the preferred diffractive technology

for optical interconnects. This follows from the greater fabricational precision and reproducible control of processes such as ion beam milling and photo-lithographic etching used to make surface-relief gratings. However, it still remains difficult to achieve efficiencies as high as volume HOEs with such gratings and during the period of research described in this thesis, DCG played an important role in many optical processing and communications projects.

Satisfying one or more of the following four criteria would make optically recorded volume HOEs more competitive with surface-relief gratings:

1. Highly accurate (nanometre resolution) micro-stepping stages to translate optical components and holographic plates.

2. Greater stability of laser output during the step and repeat processes required to optically record large arrays (greater than 16×16) of volume interconnects. Single mode diode lasers with powers in excess of 1W (across the visible and IR spectrum) would prove better than traditional gas lasers.

3. The ability to couple all of the single mode laser output power into optical fibres (perhaps by pigtailing) to facilitate simple and accurate positioning of recording sources with respect to the holographic plate; and

4. Volume media that can be sensitised to bandwidths across the visible and near IR wavelength spectrum (eliminating the need for wavelength-shifting techniques) and which do not change physical characteristics during processing, (DuPont photopolymer approaches this ideal).

Despite the difficulty of mass producing DCG HOEs and its reputation as being difficult to work with, it is still invaluable as a medium for developing new interconnect ideas, where no other technology can currently meet the requirements, e.g. high efficiency SVIs. In addition, the theoretical work and many of the recording techniques described in this thesis are equally applicable to the newer photopolymers which may be capable of implementing efficient interconnects more accurately and reproducibly than DCG.

Appendix A

Issues relating to approximate grating theories

In section 3.1 the discussion of approximate theoretical models for grating diffraction makes reference to two specific issues, (not relevant to rigorous theories):

- Optical thickness; and
- The validity of approximate theories.

This appendix briefly addresses these subjects, drawing attention to references where a more detailed analysis can be found.

A.1 Optical thickness of holograms

The optical thickness of holograms is important from the point of view of certain approximate theories that make assumptions about the 'thickness' of a grating. Therefore they can only be applied to gratings falling into the appropriate thickness regime. There are two distinct thickness regimes: 'thick' and 'thin'. However, the boundary between these two regimes is indistinct leading to the definition of an intermediate-regime. This subject has been tackled by many authors [KOG69, BEN80, MOH80a, GAY85] and [HAR84, pages 58-60], all have a slightly different definition of the key parameters used to identify the regimes, but all lead to similar

conclusions. This appendix adopts the convention of Gaylord & Moharam [GAY85]; the two main classifications are as follows:

1. **Thin** gratings. Exhibit very little angular or wavelength selectivity *OR* can be appropriately modelled by the Raman-Nath diffraction theory [RAM35, RAM36]. As a consequence of the latter definition they are termed Raman-Nath regime gratings; most surface-relief gratings are 'thin'.

2. **Thick** gratings. Exhibit a strong angular and wavelength selectivity *OR* can be appropriately modelled by the two-wave, Kogelnik diffraction theory [KOG69]. The strong Bragg-diffraction effects of these gratings classifies them as Bragg regime gratings.

With reference to section 3.2, Gaylord & Moharam define the following parameters for TE polarised light:

$$Q' = \frac{Q}{\cos\theta} = \frac{2\pi\lambda d}{n_2\Lambda^2\cos\theta}, \tag{A.1}$$

$$\gamma = \frac{\pi\Delta n_2 d}{2n_2\lambda\cos\theta}\ ;\ \text{and} \tag{A.2}$$

$$\rho = \frac{Q'}{2\gamma} \tag{A.3}$$

For TM polarised light $\gamma = \gamma\cos 2\theta$. The definitions of 'thick' and 'thin' gratings according to the relevant diffraction regime can now be quantified in terms of these parameters. However, the definitions according to the selectivity of the grating are specified in terms of the ratio d/Λ. These definitions are summarised in table A.1, any grating falling between the Raman-Nath and Bragg regimes is deemed to lie in the intermediate regime and cannot be adequately modelled by an approximate theory.

Grating type	Diffraction regime	Angular and wavelength selectivity
Thin	$Q'\gamma \leq 1$ & $Q' \ll 1$	$d/\Lambda < 10$
Thick	$\rho \geq 10$ & $Q' \gg 1$	$d/\Lambda > 10$

Table A.1: Criteria for thick and thin gratings

A.2 Validity of approximate volume grating theories

The validity of the approximate theories discussed in chapter 3 has been investigated by Moharam & Gaylord [MOH81b, MOH83b]. A generalisation of the requirements that make an approximate theory valid for analysis of a particular type of grating can be summarised as follows:

- **Transmission gratings:** TE and TM polarisations require the inclusion of higher-order diffracted waves, but do not need second-order derivatives or boundary diffraction conditions.

- **Reflection gratings:** TM polarisation requires the inclusion of higher-order diffracted waves, second-order derivatives and boundary diffraction conditions, (completely rigorous analysis). Whilst TE polarisation can ignore higher-order diffracted waves.

Appendix B

Fringe structure of HOEs with focusing power

To understand how to record holographic lenses and compensate for the aberrations induced by a wavelength shift it is helpful to understand what form the holographic fringes take through the depth of the medium and on the surface. Indeed, the recording method described in section 4.3, is based on the idea that the surface fringe structure is similar to that of a Fresnel zone plate; this is shown to be the case in this appendix. To calculate these profiles, consider the time independent interference of two spherical waves emanating from point sources R and S, as illustrated in figure B.1. The $y-$component of the electric field, E_y, at a point P, distance r from the origin, inside the holographic medium can be expressed as the superposition of the two spherical waves:

$$E_y(r) = A_R(r)\exp i[kd_1(r) + \delta_1] + A_S(r)\exp i[kd_2(r) + \delta_2]\,, \qquad \text{(B.1)}$$

where $k = {}^{2\pi n}/_{\lambda}$, A_R and A_S are the amplitudes of the wavefronts emanating from R and S respectively and δ_1 and δ_2 are the phases of the two waves at R and S respectively. The holographic fringes are described by the intensity distribution of the electric field given by

$$I(r) \;=\; E_y(r)E_y^*(r)$$

$$
\begin{aligned}
&= A_R^2(r) + A_S^2(r) \\
&\quad + A_R(r)A_S(r)\exp i\{k[d_1(r) - d_2(r)] + [\delta_1 - \delta_2]\} \\
&\quad + A_R(r)A_S(r)\exp -i\{k[d_1(r) - d_2(r)] + [\delta_1 - \delta_2]\} \\
&= A_R^2(r) + A_S^2(r) \\
&\quad + 2A_R(r)A_S(r)\cos\{k[d_1(r) - d_2(r)] + \delta\}\,, \qquad \text{(B.2)}
\end{aligned}
$$

where $\delta = \delta_1 - \delta_2$. The loci of constructive and destructive interference correspond to the maxima and minima of the cosine function, respectively. So that:

$$d_1(r) - d_2(r) = \frac{m\lambda}{2n} - \delta = \Delta \qquad \text{(B.3)}$$

With reference to figure B.1, where R is a distance f_0 from the origin and at an

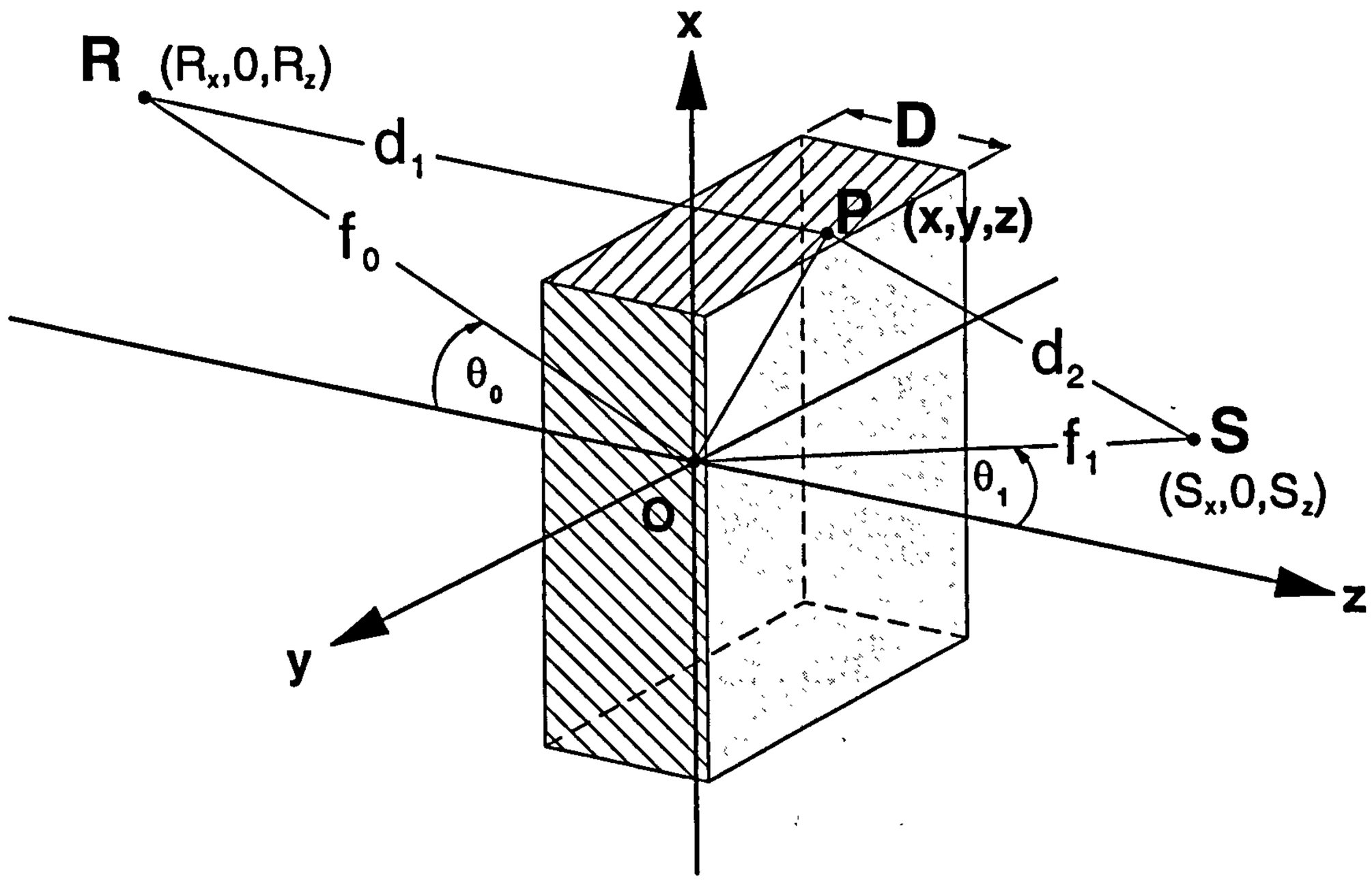


Figure B.1: Off-axis interference geometry for two point sources.

angle θ_0 to the $z-$axis and S is at a distance f_1 from the origin and at an angle θ_1 to the $z-$axis:

$$d_1(r) \;=\; \sqrt{(x + f_0\sin\theta_0)^2 + y^2 + (z + f_0\cos\theta_0)^2}\;;\ \text{and}$$

$$d_2(r) \;=\; \sqrt{(f_1 \sin\theta_1 - x)^2 + y^2 + (f_1 \cos\theta_1 - z)^2} \tag{B.4}$$

Substituting equations B.4 into equation B.3 results in a quadratic equation, in x, y and z, of the form:

$$Ax^2 + By^2 + Cz^2 + Dxz + Ex + Fz + G = 0\ , \tag{B.5}$$

where

$$\begin{aligned}
A &= 4\left[\Delta^2 - (f_0 \sin\theta_0 + f_1 \sin\theta_1)^2\right]\ , \\
B &= 4\left[\Delta^2\right]\ , \\
C &= 4\left[\Delta^2 - (f_0 \cos\theta_0 + f_1 \cos\theta_1)^2\right]\ , \\
D &= -8(f_0 \sin\theta_0 + f_1 \sin\theta_1)(f_0 \cos\theta_0 + f_1 \cos\theta_1)\ , \\
E &= 4\left[(\Delta^2 + f_1^2 - f_0^2)(f_0 \sin\theta_0 + f_1 \sin\theta_1) - 2\Delta^2 f_1 \sin\theta_1\right]\ , \\
F &= 4\left[(\Delta^2 + f_1^2 - f_0^2)(f_0 \cos\theta_0 + f_1 \cos\theta_1) - 2\Delta^2 f_1 \cos\theta_1\right]\ ; \text{ and} \\
G &= 4\Delta^2 f_1^2 - (f_0^2 - f_1^2 - \Delta^2)^2
\end{aligned} \tag{B.6}$$

By analysing this quadratic equation in both the $x-y$ plane and the $x-z$ plane, the fringe profile at the surface and through the depth of the medium can be found.

Depth profile

The fringes can be examined in the $x-z$ plane with no loss of generality. Looking in the $x-z$ plane, by setting $y=0$, the quadratic takes the form of a family of hyperbolas with foci governed by the constants D, E and F in equation B.6 above. This has one exception; in the case of the source of the wavefront R being at infinity (a plane wave) and on-axis ($\theta_0 = 0$), in which case the fringes are parabolic, (this is not immediately obvious from equation B.5, for more detail refer to [STO86, pages 66-74]). The fringe structures will, of course, be far from this ideal in a real holographic medium, such as DCG, where nonuniformities will modify their form.

Surface profile

The surface profile of the holographic fringes is of much more interest with regard to HOEs with focusing power, since it is the surface diffraction which determines the diffracted direction [CLO75, page 409]. Setting $z = 0$ restricts us to examining the $x - y$ plane, the quadratic then takes the form of a family of ellipses, (the major and minor axes of which are equal to $\sqrt{B}$ and $\sqrt{A}$ respectively), with foci on the $x-$axis. The exceptions in this case are when R and S are on-axis, (θ_0 and θ_1 are both zero), and when the wavefront coming from R is plane, (on- or off-axis), in both instances the surface fringe profiles are circular, (again for on-axis detail see [STO86]). Generally the fringe structure is elliptical and closely resembles a thin Fresnel zone plate, (see [HEC87, page 445]) in appearance; this is utilised to determine where the ideal recording points for a wavelength-shifted lenslet are located in space, in section 4.3.

Appendix C

The applicability of the constant brightness theorem to optical interconnects

Light cannot be fanned-in and out to a given volume in space in an arbitrary fashion without constraining the power in the interconnected beams (optical communication channels) - there is a limit to what can be achieved. Several authors have considered this limit, determined by the constant brightness (or radiance) theorem [MAR72, BOY83, GOO85, FRI86, WEL89]. Although it is an important consideration in optical communication theory and a generally understood and accepted theorem of radiometry, it is clarified here with regard to the types of interconnects described in this thesis.

C.1 The constant brightness/radiance theorem

Otherwise known as *the principle of generalised etendue* or *Lagrange invariant*, the constant-radiance theorem is a theorem that can be derived from geometric optics or from thermodynamics [BOY83, section 5.2], [DRU59, pages 502-505] as it is a direct corollary of the Clausius statement of the second law. In its simplest form it states that:

No passive linear optical system can increase the radiance ($Wm^{-2}sr^{-1}$) *of an optical beam.*

If a beam has a cross-sectional area A and occupies a solid angle Ω then it can be expressed mathematically as:

$$A \times \Omega \; = \; etendue \; = \; \text{constant} \; , \tag{C.1}$$

throughout the system.

The theorem is analogous to Liouville's theorem of statistical mechanics [MAR72, pages 112-124], which when applied to wave optics can be stated as:

The volume in phase space, filled with representation points of a bundle of rays, remains constant.

The radiance theorem is often, for convenience, referred to as Liouville's theorem, but there is a fundamental difference in meaning. This arises from the fact that Liouville's theorem makes a statistical statement about the average density of points in phase space, whereas the radiance theorem is deterministic [WEL89, appendix A].

C.2 Implications of the radiance theorem

The implications of the radiance theorem to *coherent* and *incoherent* optical interconnection systems are not immediately apparent. The theorem governs the amount of power that can be expected in a single spatial mode from an N-fold fan-out or fan-in based upon the radiance being constant throughout any given optical system and the etendue (defined by equation C.1) being constant along each optical path within that system.

C.2.1 Fan-out

The ramifications with respect to fan-out are fairly intuitive for both coherent and incoherent light. If the power in an optical beam is divided equally N ways into beams of the same A and Ω, then the radiance in each beam is $1/N^{th}$ of the input.

Alternatively, since Ω is proportional to the square of the numerical aperture of the beam, irradiance (Wm^{-2}) will be constant between the input and an output beam as long as the numerical aperture of each output beam is increased by a factor of $\sqrt{N}$, (see figure C.1a).

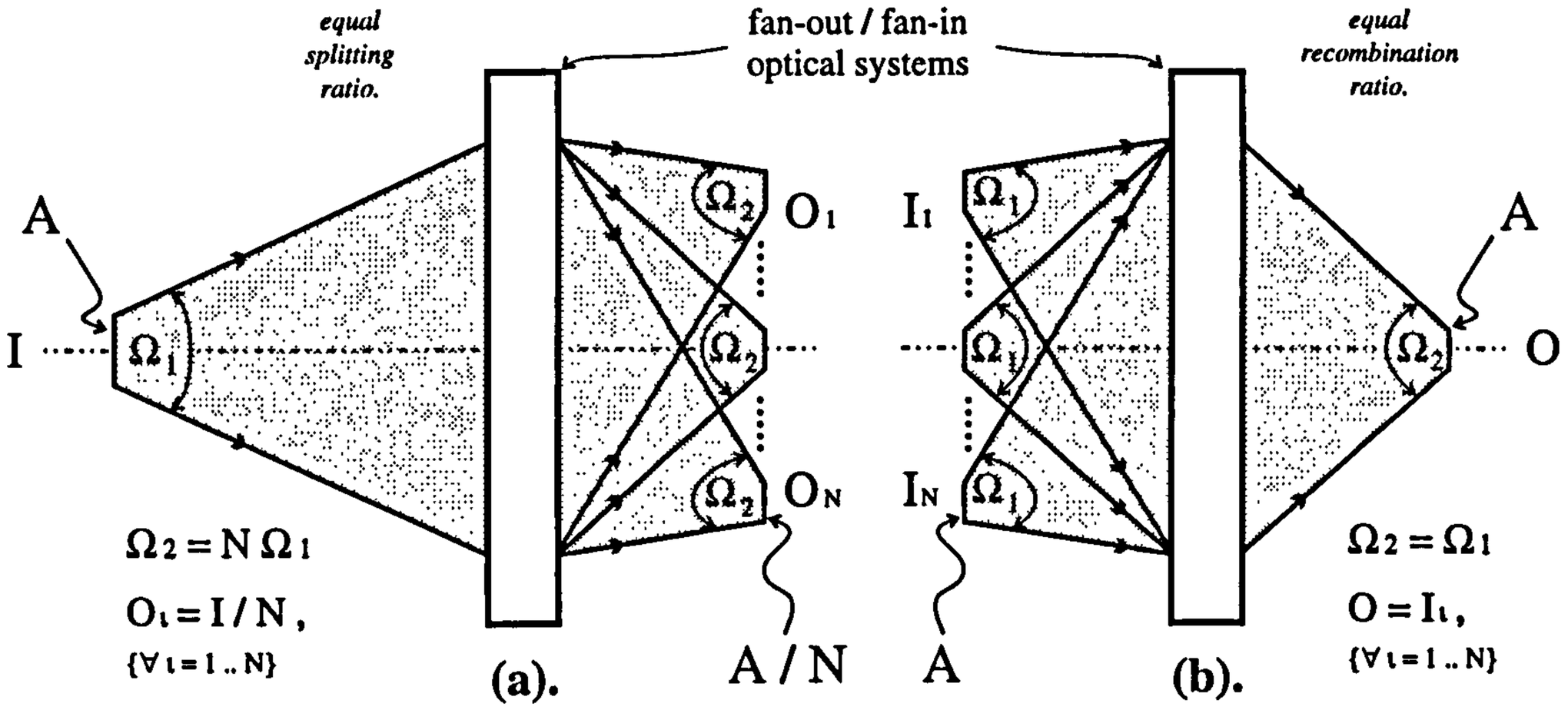


Figure C.1: Conditions for constant irradiance (Wm^{-2}) between all input, I, and output, O beams, for: (a). Fan-out from; and (b).Fan-in to, the same spatial mode.

For an optical system of fixed f/number, an N-fold fan-out will result in an N-fold irradiance reduction in each interconnection.

C.2.2 Fan-in

The implications with regard to fan-in are less obvious. The two distinct cases of sets of beams within optical interconnection systems which are *mutually coherent* (derived from the same source) and *mutually incoherent*, (uncorrelated coherent sources or, alternatively, completely incoherent light), must be considered separately:

Incoherent fan-in

If N identical beams of cross-sectional area A_1, subtending a solid angle Ω_1 are fanned-in to form a single beam of area A_2, subtending a solid angle Ω_2 then the maximum fraction of the total optical power (from the N beams) in the resultant beam (since radiance must be constant) will be given by $\Omega_2 A_2 / N \Omega_1 A_1$. Thus if

N identical beams fan-in to a single beam of the same divergence and cross-sectional area, the optical power in the resultant beam is, at most, $1/N^{th}$ of that in all of the beams. Alternatively, irradiance (Wm^{-2}) will be constant between the resultant and an incident beam when the numerical aperture of the output beam is equal to that of the input, (see figure C.1b).

Consequently, for an optical system of fixed f/number, an N-fold fan-in will result in the same irradiance at the output as at each input. To understand the physical reasons why this must be the case we cannot purely think about beams combining, consideration must be given as to *how* multiple beams are made to fan-in to a single beam (or spatial mode).

Figure C.2 shows four typical ways in which an optical interconnection system can combine two beams into one. When using these systems with incoherent light the following considerations apply:

- **Aperture division:** Only half of the available aperture of the lens (focal length f) is being utilised for any one beam. Consequently, although the resultant beam, O, would appear to have the same NA as I_1 and I_2 and contains all of the input power, the focal spot will be larger because the effective NA (for each beam) is halved, (note that this does not fan-in to the same spatial mode). The radiance will be constant throughout the system, so that the irradiance at O is, at most, 50% of $I_1 + I_2$.

- **Beam splitter:** For a fixed ratio non-polarising beam splitter the maximum power in either of the two output directions is 50% of the total input.

- **HOE:** If the HOE is replayed on-Bragg in the backwards travelling (fan-out) direction by the O-beam, then at most 50% of the incident light can be sent into each of I_1 and I_2. Using the principle of optical reversibility, each grating can be, at most, 50% efficient and so a maximum of 50% of the input power ($I_1 + I_2$) can reach O when used in the fan-in configuration shown; and

- **Y-junction waveguide:** The same reversibility argument used on the HOE applies with this type of waveguide combiner [JAC95]. In accordance with

the radiance theorem, at most, only 50% of the incident irradiance ends up in the resultant direction. For single mode guides the deterministic form of the brightness theorem (as derived from geometric optics) cannot be assumed [GOO85, page 1492].

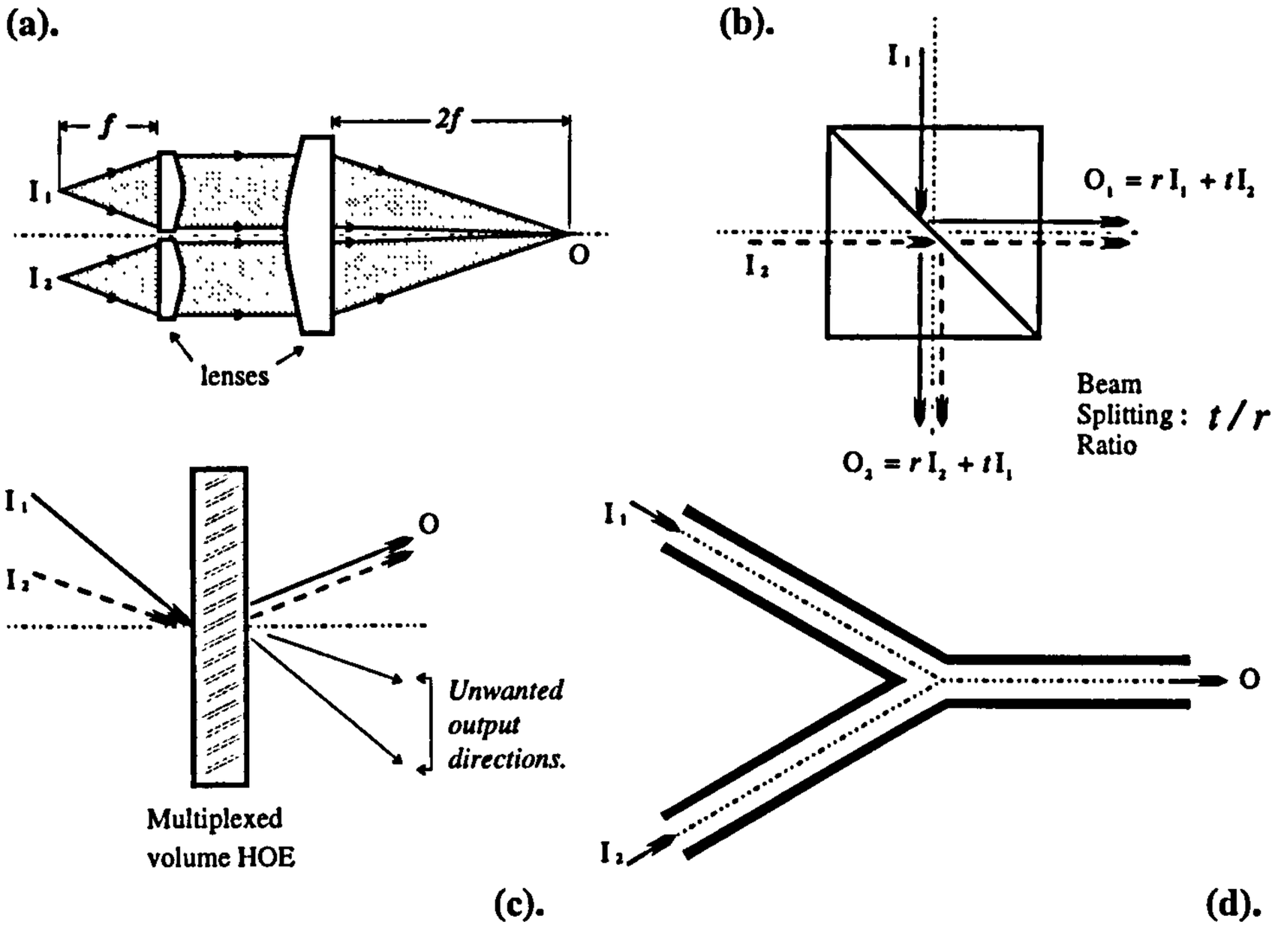


Figure C.2: Practical ways to fan-in two beams of light to a single resultant beam, (coherently or incoherently): (a). Aperture division [WAL86], (b). Beam Splitter, (for uniform coherent beams: if $t/r = 50/50$ and beams are in phase, $O_1 = 2I$), (c). HOE, (multiplexed gratings with equal efficiency, as per chapter 5); and (d). Y-junction waveguide, (single or multi-mode).

When uncorrelated coherent beams are considered, (the relative phases will be a rapidly varying function of time), the same considerations apply; at most 50% of the incident power ends up in the resultant direction, the average being over all possible phases of the incident beams.

Coherent fan-in

The same reasoning applies to the fan-in of mutually coherent beams unless the optical system is stabilised to maintain absolutely constant paths in all interconnections. The power contained in the fanned-in beam will depend on the relative

phases of the incident beams and so a statistical theorem, (that of Liouville), is more appropriate when considering coherent fan-in.

For any one instance of the relative phases between beams, the optical power delivered to the resultant beam is influenced by interference between all beams and can be more or less than that predicted by the radiance theorem. However, when averaged over all possible phases, the constant-radiance theorem implies that the resultant beam can contain no more than $1/N^{th}$ of the radiance for N-fold fan-in.

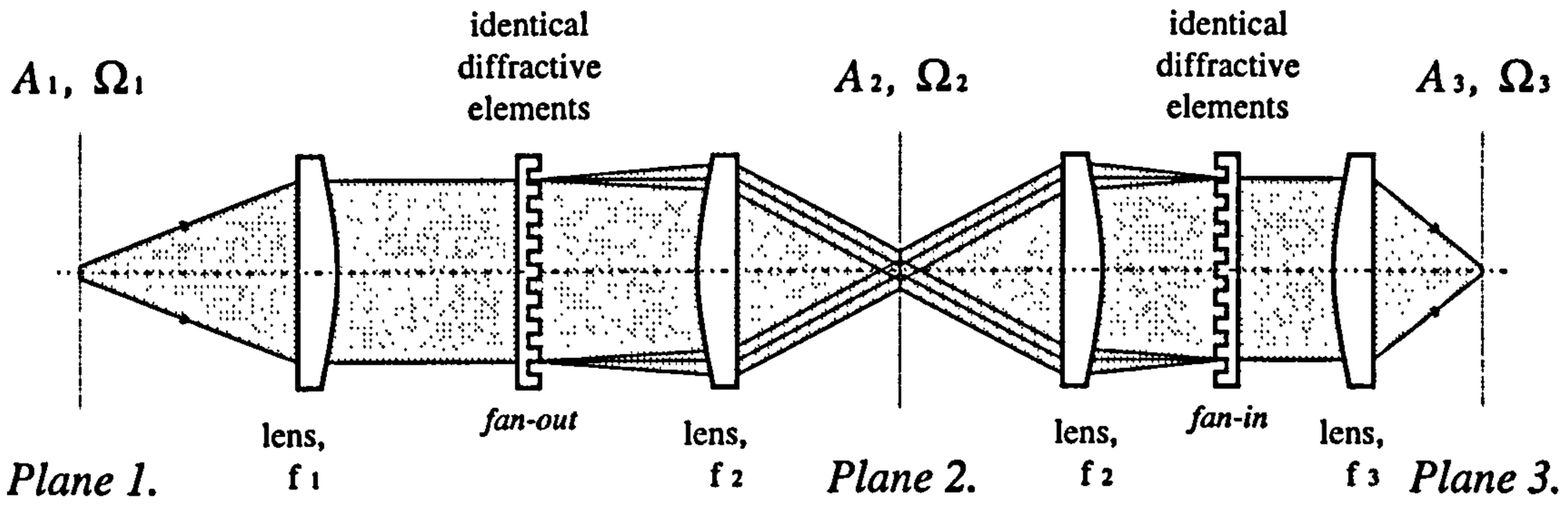


Figure C.3: Idealised optical system to create multiple mutually coherent sources, at plane 2 (each with beam cross-sectional area A_2 and solid angle Ω_2), from a single thermal (lasing) source at plane 1 (A_1, Ω_1) and fan these into a single resultant spatial mode at plane 3 (A_3, Ω_3). All phases must be preserved, geometric aberrations eliminated and path lengths stabilised for this to work, (each type of duplicate element: similar lenses and HOEs, must be identical).

Considering the beam combining methods shown in figure C.2b,c&d, if the two identical fan-in beams are in phase all of the incident power can be coupled into the outgoing direction. If this is not the case, power can be coupled into other output directions (or radiative modes for single mode waveguides [JAC95]) and interference fringe structure will be evident in the beams. Again, averaged over all possible phases, on average only 50% of the incident irradiance ends up in the required direction.

It is, however, impractical to consider two or more *separate* and distinct sources as being constantly in phase with one another since they cannot have the same thermal characteristics. If several mutually coherent beams are to fan-in to a single resultant direction (and apparently contravene the radiance theorem, when considering a single fixed phase relationship, provided the optical paths can be stabilised), they

must have originated from the same source. This is illustrated in figure C.3, where plane 2 represents the multiple sources produced by fanning out a single source using an idealised 100% efficient uniform diffractive element (e.g. a kinoform [LES69]). Consequently, when considering all paths in the entire system the radiance theorem is obeyed: nothing more than one-to-one imaging is achieved, since $A_1\Omega_1 = A_3\Omega_3$.

If time varying amplitude or phase information is imposed upon mutually coherent interconnection beams *e.g.* in an optical communication or processing system, then the phase difference between beams will, effectively, vary randomly and uniformly over 2π radians. In such a situation, although sources may be mutually coherent, the optical communication channels become mutually incoherent and the predictions of the constant-radiance theorem are maintained.

Goodman [GOO85] discusses the implications of the radiance theorem to optical interconnections in more detail, drawing parallels with electronic counterparts. Ways in which the fan-in limitations of the theorem can be circumvented are discussed in section 5.1 and by McCormick [MCC93, pages 34-36].

Appendix D

Example of the discriminatory wavevector coupling algorithm for fan-out elements

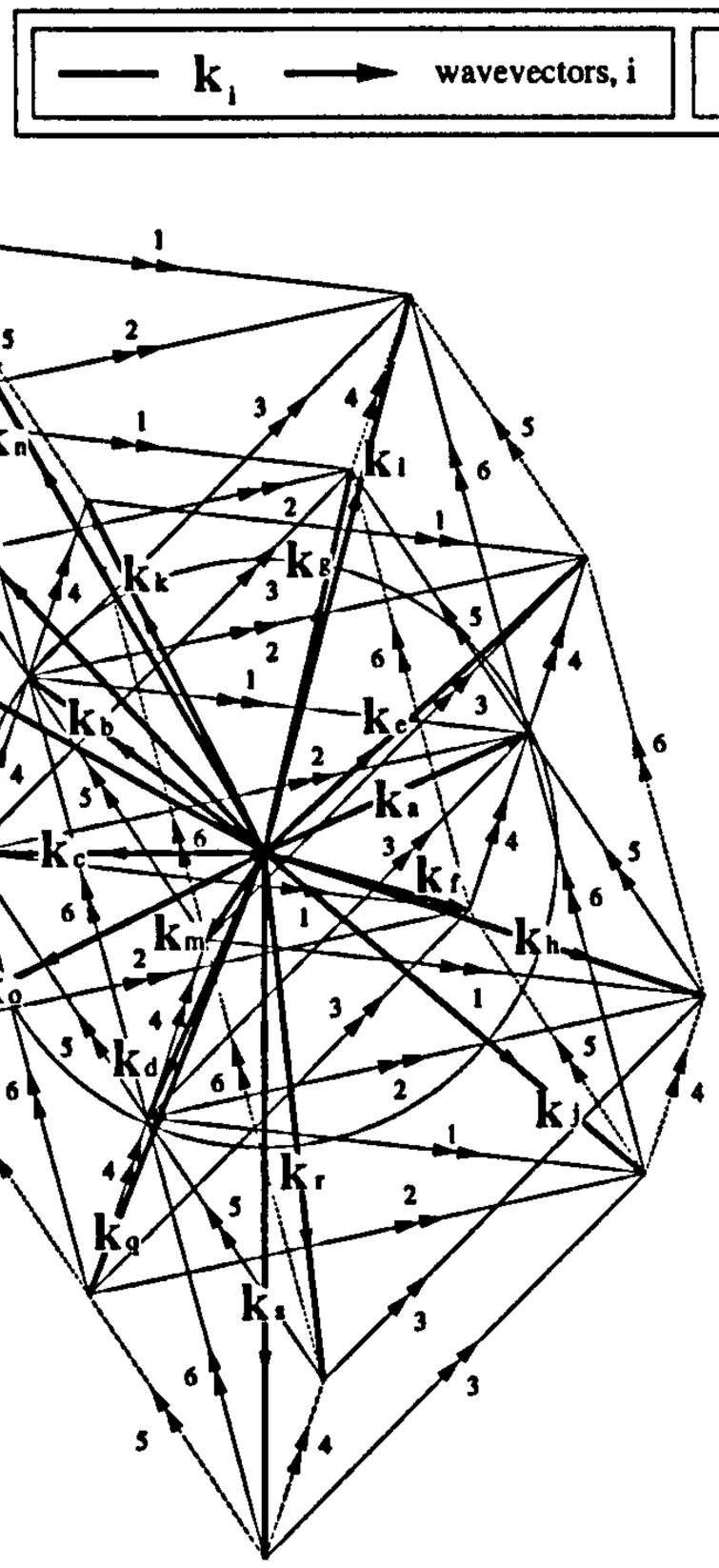
This appendix illustrates how the discriminatory algorithm, described in section 5.4.5, generates the wavevectors and the coupling between them required to model a fan-out to three grating.

D.1 Fan-out to three example

Describing the discriminatory algorithm in detail is difficult but is demonstrated by taking an example: in this case a reflection fan-out to three hologram. The vector Floquet diagram for such a grating structure is shown in figure D.1.

Three primary gratings ($\mathbf{K_1}$, $\mathbf{K_2}$ & $\mathbf{K_3}$) recorded coherently with a common reference wave will generate three secondary gratings ($\mathbf{K_4}$, $\mathbf{K_5}$ & $\mathbf{K_6}$). When replayed on-Bragg with this reference wave, $\mathbf{k_0}$, the following wavevectors will be generated under constraint 5.19:

- Primary waves: $\mathbf{k_{1,2,3}}$; the original object wave directions.
- Intermodulation waves: $\mathbf{k_{4..9}}$; from figure D.1 it can be seen that these six wavevectors are generated by the ± 1 secondary grating orders. I-waves are



K_{label}	k_m	k_b	K_a	g	
a).	0	0	0	0	replay wave.
b).	1	0	1	-1	
	16	2	4	1	
	26	3	6	1	
c).	2	0	2	-1	recording (principal) waves.
	11	1	4	-1	
	24	3	5	1	
d).	3	0	3	-1	
	15	1	6	-1	
	19	2	5	-1	
e).	4	0	4	1	Intermodulation waves (or I-waves).
f).	5	0	4	-1	
g).	6	0	5	1	
h).	7	0	5	-1	
i).	8	0	6	1	
j).	9	0	6	-1	
k).	10	1	4	1	Secondary waves (or S-waves).
l).	12	1	5	1	
	20	2	6	1	
m).	13	1	5	-1	
	22	3	4	1	
n).	14	1	6	1	
o).	17	2	4	-1	
p).	18	2	5	1	
q).	21	2	6	-1	
	23	3	4	-1	
r).	25	3	5	-1	
s).	27	3	6	-1	

Figure D.1: Reflection fan-out to three wavevector and grating coupling structure. The wave-vector descriptions (shown in the table insert [table D.1]) are derived from the vector Floquet condition, equation 5.15. Those $\mathbf{k}_m$ grouped under the same $\mathbf{k}_{label}$ are degenerate and those couplings shown dotted can be ignored under constraint 5.19.

never degenerate with other space-harmonics; and

- Secondary waves: $\mathbf{k}_{11..27}$; from the table in figure D.1 it can be seen that some secondary waves are degenerate in their descriptions with primary and other secondary waves.

The table in figure D.1 describes how each wavevector, $\mathbf{k}_m$, is generated from which root wavevector, $\mathbf{k}_b$, via which primary or secondary grating, $\mathbf{K}_a$, and from which grating order, g, according to the vector Floquet condition 5.15.

To automatically generate wavevectors and coupling, a table such as that in figure D.1 is compiled. Contained in this table of wavevector descriptions (that includes degenerate descriptions) is all the coupling information shown in the wavevector di-

agram of the same figure. This information is extracted to form a coupling matrix for *all* wavevectors, $\mathbf{k}_m$, (where $\mathbf{k}_i, \mathbf{k}_j \, \epsilon \, \mathbf{k}_m$ and $j < i \leq m_{max}$), using the following rules:

1. If the $\mathbf{k}_b$ of wavevector $\mathbf{k}_i$ is equal to $\mathbf{k}_j$, then $\mathbf{k}_i$ couples to $\mathbf{k}_j$ via the $\mathbf{K}_a$ of $\mathbf{k}_i$.

2. If $\mathbf{k}_i$ and $\mathbf{k}_j$ have the same $\mathbf{K}_a$ and the same g, then $\mathbf{k}_i$ couples to $\mathbf{k}_j$ via the $\mathbf{K}_a$ of $\mathbf{k}_i$.

3. If $\mathbf{k}_i$ and $\mathbf{k}_j$ have the same $\mathbf{K}_a$, the same g and if the $\mathbf{k}_b$ of $\mathbf{k}_i$ and $\mathbf{k}_j$ are the recording beam wavevectors forming a secondary *or* primary grating $\mathbf{K}_R$, then $\mathbf{k}_i$ couples to $\mathbf{k}_j$ via $\mathbf{K}_R$; and

4. If $\mathbf{k}_i$ and $\mathbf{k}_j$ have the same $\mathbf{k}_b$ and the $\mathbf{K}_a$ of $\mathbf{k}_i$ and $\mathbf{k}_j$ are the primary gratings that close a vector triangle with secondary grating $\mathbf{K}_S$, then $\mathbf{k}_i$ couples to $\mathbf{k}_j$ via $\mathbf{K}_S$.

The constraint 5.19 eliminates the need for some of the coupling illustrated in figure D.1. In the same way that wavevector descriptions are only generated for -1 primary and ± 1 secondary grating interactions, coupling between these wavevectors is limited to a maximum of two grating interactions from $\mathbf{k}_0$, (provided that one of them is via a primary grating otherwise only a single interaction is allowed). This leads to no secondary grating coupling between I-waves or distinct S-waves, whilst maintaining primary grating coupling between I-waves and both S-waves and primary waves. Consequently, the couplings shown dotted in figure D.1 are not included in the analysis.

Once the coupling matrix is established, degenerate descriptions are removed from the analysis and the coupling matrix altered accordingly. The degenerate wavevectors are identified by examining the $x-$direction wavevector components, β_m, (from equation 5.16) and comparing them, for all space-harmonics, to an accuracy of 10^{-14}.

Table D.2 shows a representation of the coupling matrix for the fan-out to three case when the coupling rules (described above) are applied, (secondary degenerate

D.1 Fan-out to three example

k	0	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27
0	ζ_0	χ_1	χ_2	χ_3	χ_4	χ_4	χ_5	χ_5	χ_6	χ_6	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
1	χ_1	ζ_1	χ_4	χ_6	0	0	0	0	0	0	χ_4	χ_4	χ_5	χ_5	χ_6	χ_6	0	0	0	0	0	0	0	0	0	0	0	0
2	χ_2	χ_4	ζ_2	χ_5	0	0	0	0	0	0	0	0	0	0	0	0	χ_4	χ_4	χ_5	χ_5	χ_6	χ_6	0	0	0	0	0	0
3	χ_3	χ_6	χ_5	ζ_3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	χ_4	χ_4	χ_5	χ_5	χ_6	χ_6
4	χ_4	0	0	0	ζ_4	0	0	0	0	0	χ_1	0	0	0	0	0	χ_2	0	0	0	0	0	χ_3	0	0	0	0	0
5	χ_4	0	0	0	0	ζ_5	0	0	0	0	0	χ_1	0	0	0	0	0	χ_2	0	0	0	0	0	χ_3	0	0	0	0
6	χ_5	0	0	0	0	0	ζ_6	0	0	0	0	0	χ_1	0	0	0	0	0	χ_2	0	0	0	0	0	χ_3	0	0	0
7	χ_5	0	0	0	0	0	0	ζ_7	0	0	0	0	0	χ_1	0	0	0	0	0	χ_2	0	0	0	0	0	χ_3	0	0
8	χ_6	0	0	0	0	0	0	0	ζ_8	0	0	0	0	0	χ_1	0	0	0	0	0	χ_2	0	0	0	0	0	χ_3	0
9	χ_6	0	0	0	0	0	0	0	0	ζ_9	0	0	0	0	0	χ_1	0	0	0	0	0	χ_2	0	0	0	0	0	χ_3
10	0	χ_4	0	0	χ_1	0	0	0	0	0	ζ_{10}	0	0	0	0	0	χ_4	0	0	0	0	0	χ_6	0	0	0	0	0
11	0	χ_4	0	0	0	χ_1	0	0	0	0	0	ζ_{11}	0	0	0	0	0	χ_4	0	0	0	0	0	χ_6	0	0	0	0
12	0	χ_5	0	0	0	0	χ_1	0	0	0	0	0	ζ_{12}	0	0	0	0	0	χ_4	0	0	0	0	0	χ_6	0	0	0
13	0	χ_5	0	0	0	0	0	χ_1	0	0	0	0	0	ζ_{13}	0	0	0	0	0	χ_4	0	0	0	0	0	χ_6	0	0
14	0	χ_6	0	0	0	0	0	0	χ_1	0	0	0	0	0	ζ_{14}	0	0	0	0	0	χ_4	0	0	0	0	0	χ_6	0
15	0	χ_6	0	0	0	0	0	0	0	χ_1	0	0	0	0	0	ζ_{15}	0	0	0	0	0	χ_4	0	0	0	0	0	χ_6
16	0	0	χ_4	0	χ_2	0	0	0	0	0	χ_4	0	0	0	0	0	ζ_{16}	0	0	0	0	0	χ_5	0	0	0	0	0
17	0	0	χ_4	0	0	χ_2	0	0	0	0	0	χ_4	0	0	0	0	0	ζ_{17}	0	0	0	0	0	χ_5	0	0	0	0
18	0	0	χ_5	0	0	0	χ_2	0	0	0	0	0	χ_4	0	0	0	0	0	ζ_{18}	0	0	0	0	0	χ_5	0	0	0
19	0	0	χ_5	0	0	0	0	χ_2	0	0	0	0	0	χ_4	0	0	0	0	0	ζ_{19}	0	0	0	0	0	χ_5	0	0
20	0	0	χ_6	0	0	0	0	0	χ_2	0	0	0	0	0	χ_4	0	0	0	0	0	ζ_{20}	0	0	0	0	0	χ_5	0
21	0	0	χ_6	0	0	0	0	0	0	χ_2	0	0	0	0	0	χ_4	0	0	0	0	0	ζ_{21}	0	0	0	0	0	χ_5
22	0	0	0	χ_4	χ_3	0	0	0	0	0	χ_6	0	0	0	0	0	χ_3	0	0	0	0	0	ζ_{22}	0	0	0	0	0
23	0	0	0	χ_4	0	χ_3	0	0	0	0	0	χ_6	0	0	0	0	0	χ_5	0	0	0	0	0	ζ_{23}	0	0	0	0
24	0	0	0	χ_5	0	0	χ_3	0	0	0	0	0	χ_6	0	0	0	0	0	χ_5	0	0	0	0	0	ζ_{24}	0	0	0
25	0	0	0	χ_5	0	0	0	χ_3	0	0	0	0	0	χ_6	0	0	0	0	0	χ_5	0	0	0	0	0	ζ_{25}	0	0
26	0	0	0	χ_6	0	0	0	0	χ_3	0	0	0	0	0	χ_6	0	0	0	0	0	χ_5	0	0	0	0	0	ζ_{26}	0
27	0	0	0	χ_6	0	0	0	0	0	χ_3	0	0	0	0	0	χ_6	0	0	0	0	0	χ_5	0	0	0	0	0	ζ_{27}

Table D.2: Tabular representation of the coupling matrix for fan-out to three with all space-harmonics included, (index **k** represents the wavevectors shown in figure D.1). This is not the true coupling matrix, but is an intermediate stage in deriving the final form, (the highlighted areas are the degenerate orders).

k	0	1	2	3	4	5	6	7	8	9	10	12	13	14	17	18	21	25	27
0	ζ_0	χ_1	χ_2	χ_3	χ_4	χ_4	χ_5	χ_5	χ_6	χ_6	0	0	0	0	0	0	0	0	0
1	χ_1	ζ_1	χ_4	χ_6	χ_2	0	0	0	χ_3	0	χ_4	χ_5	χ_5	χ_6	0	0	0	0	0
2	χ_2	χ_4	ζ_2	χ_5	0	χ_1	χ_3	0	0	0	0	χ_6	0	0	χ_4	χ_5	χ_6	0	0
3	χ_3	χ_6	χ_5	ζ_3	0	0	0	χ_2	0	χ_1	0	0	χ_4	0	0	0	χ_4	χ_5	χ_6
4	χ_4	χ_2	0	0	ζ_4	0	0	0	0	0	χ_1	0	χ_3	0	0	0	0	0	0
5	χ_4	0	χ_1	0	0	ζ_5	0	0	0	0	0	0	0	0	χ_2	0	χ_3	0	0
6	χ_5	0	χ_3	0	0	0	ζ_6	0	0	0	0	χ_1	0	0	0	χ_2	0	0	0
7	χ_5	0	0	χ_2	0	0	0	ζ_7	0	0	0	0	χ_1	0	0	0	0	χ_3	0
8	χ_6	χ_3	0	0	0	0	0	0	ζ_8	0	0	χ_2	0	χ_1	0	0	0	0	0
9	χ_6	0	0	χ_1	0	0	0	0	0	ζ_9	0	0	0	0	0	0	χ_2	0	χ_3
10	0	χ_4	0	0	χ_1	0	0	0	0	0	ζ_{10}	0	χ_6	0	0	0	0	0	0
12	0	χ_5	χ_6	0	0	0	χ_1	0	χ_2	0	0	ζ_{12}	0	χ_4	0	χ_4	0	0	0
13	0	χ_5	0	χ_4	χ_3	0	0	χ_1	0	0	χ_6	0	ζ_{13}	0	0	0	0	χ_6	0
14	0	χ_6	0	0	0	0	0	0	χ_1	0	0	χ_4	0	ζ_{14}	0	0	0	0	0
17	0	0	χ_4	0	0	χ_2	0	0	0	0	0	0	0	0	ζ_{17}	0	χ_5	0	0
18	0	0	χ_5	0	0	0	χ_2	0	0	0	0	χ_4	0	0	0	ζ_{18}	0	0	0
21	0	0	χ_6	χ_4	0	χ_3	0	0	0	χ_2	0	0	0	0	χ_5	0	ζ_{21}	0	χ_5
25	0	0	0	χ_5	0	0	0	χ_3	0	0	0	0	χ_6	0	0	0	0	ζ_{25}	0
27	0	0	0	χ_6	0	0	0	0	0	χ_3	0	0	0	0	0	0	χ_5	0	ζ_{27}

Table D.3: Tabular representation of the final coupling matrix for fan-out to three with degenerate space-harmonic information removed, (index **k** represents the wavevectors shown in figure D.1).

D.1 Fan-out to three example

wavevectors are highlighted). To form the final coupling matrix the coupling information is copied from each degenerate order into the sole retained representation of that wavevector direction; e.g. the row and column information for $\mathbf{k}_{16}$ and $\mathbf{k}_{26}$ is copied into $\mathbf{k}_1$. The degenerate wavevectors are then removed from the matrix and the result is shown in table D.3, (tables D.2 & D.3 contain demarcations denoting the replay, primary, I- and S-waves).

This coupling algorithm still retains some coupling between S-waves, despite the stipulations of constraint 5.19, as illustrated in table D.3, (where $\mathbf{k}_i$ couples to $\mathbf{k}_j$ and $i, j \geq 10$). These can be set to zero or retained, it makes no significant difference to the numerical results.

Glossary

ADC	Ammonium Dichromate
AR	Anti-Reflection
ATM	Asynchronous Transmission Mode
ATT	Amplitude Transmission Theory
BEAT	Bistable Etalon with Absorbed Transmission
BER	Bit Error Rate
BKK	Burckhardt Kaspar & Knop
BPG	Binary Phase Grating
CCD	Charge Coupled Device
CGH	Computer Generated Hologram
CLIP	Cellular Logic Image Processor
CPU	Central Processing Unit
CWM	Coupled-Wave Method
CWT	Coupled-Wave Theory
DCG	Dichromated Gelatin
EGSN	Extended Generalised Shuffle Network
EM	Electro-Magnetic
FFT	Fast Fourier Transform
FLC	Ferro-electric Liquid Crystal
FWHM	Full Width Half Maximum
FZP	Fresnel Zone Plate
GRIN	Gradient-Index
HAtt	Holographic Attenuator
HOE	Holographic Optical Element
HUD	Head-Up Display
IC	Integrated Circuit
IR	Infra-Red
I-Wave	Intermodulation Wave
LAN	Local Area Network
LCLV	Liquid Crystal Light Valve
MIMD	Multiple Instruction Multiple Data
MM	Modal Method
MMFE	Modal Method with a Fourier Expansion
MMME	Modal Method with a Modal Expansion
NA	Numerical Aperture
ND	Neutral Density
Nd:YAG	Neodymium Yttrium Aluminium Garnet
NLIF	Non-Linear Interference Filter
NNI	Nearest Neighbour Interconnect
O-CLIP	Optical-Cellular Logic Image Processor

OCPM Optically Connected Parallel Machines

PBS Polarising Beam Splitter

PC Personal Computer

PMR Patterned Mirror Reflector

PRG Pure Reflection Grating

RC Resistance-Capacitance

RCWM Rigorous Coupled Wave Method

RCWT Rigorous Coupled Wave Theory

RH Relative Humidity

RISC Reduced Instruction Set Computer

RMS Root Mean Square

SBWP Space-Bandwidth Product

SEED Self Electro-optic Effect Device

SII Spatially Invariant Interconnect

SIMD Single Instruction Multiple Data-stream

SISD Single Instruction Single Data-stream

SLM Spatial Light Modulator

SNR Signal-to-Noise Ratio

S-SEED Symmetric-Self Electro-optic Effect Device

SVHOE Space-Variant HOE

SVI Spatially Variant Interconnect

S-Wave Secondary Wave

TE Transverse Electric

TEM Transverse Electro-Magnetic

TGDM Thin Grating Decomposition Method

TM Transverse Magnetic

UV Ultra-Violet

VCSEL Vertical Cavity Surface Emitting Laser

VLSI Very Large Scale Integration

WAN Wide Area Network

References

[ABR83] E. Abraham, C. T. Seaton and S. D. Smith, "The optical computer.", Scientific American, 63-71, January (1983).

[ACK94] B. Acklin and J. Jahns, "Packaging considerations for planar optical interconnection systems.", Applied Optics, **33**, 1391-1397 (1994).

[ALF75a] R. Alferness and S. K. Case, "Coupling in double exposed, thick holographic gratings.", J. Opt. Soc. Am., **65**, 730-739 (1975).

[ALF75b] R. Alferness, "Analysis of optical propagation in thick holographic gratings.", Applied Physics, **7**, 29-33 (1975).

[ALF76] R. Alferness, "Analysis of propagation at the second order Bragg angle of a thick holographic grating.", J. Opt. Soc. Am., **66**, 353-362 (1976).

[ASS88] M. Assenheimer, Y. Amitai and A. A. Friesem, "Recursive design for an efficient HOE with different recording and readout wavelengths.", Applied Optics, **27**, 4747-4752 (1988).

[BAC78] J. Backus, "Can programming be liberated from the von Neumann style? A functional style and its algebra of programs.", Communications of the ACM, **21**, 613-641, August (1978).

[BAC82] J. Backus, "Function-level computing.", IEEE Spectrum, 22-37, August (1982).

[BAR82] H. Bartelt and S. K. Case, "High-efficiency hybrid computer-generated holograms.", Applied Optics, **21**, 2886-2890 (1982).

[BAR83] H. Bartelt and S. K. Case, "Coordinate transformations via multifacet holographic optical elements.", Optical Engineering, **22**, 497-500 (1983).

[BAR95] C. P. Barrett, "Reconfigurable free-space optical interconnection systems.", Ph.D. thesis, Heriot-Watt University, U.K. (1995).

[BEN79] B. Benlarbi and L. Solymar, "The effect of the relative intensity of the reference beam on the reconstruction properties of volume phase holograms.", Optica Acta, **26**, 271-278 (1979).

[BEN80] B. Benlarbi, D. J. Cooke and L. Solymar, "Higher order modes in thick phase gratings.", Optica Acta, **27**, 885-895 (1980).

[BER67] L. Bergstein and D. Kermisch, "Image storage and reconstruction in volume holography.", Proc. Symp. on Modern Optics, **17**, 655-680 (1967).

[BIE77] K. Biedermann, *Silver halide photographic materials* chapter 2 in "Holographic recording materials.", Topics in Applied Physics, **20**, *editor*: H. M. Smith, Springer-Verlag, Berlin, Germany (1977).

[BIR91] K. D. Bird, D. Daly and T. J. Hall, "Computer generated microlens arrays and their application to optical free space switching networks.", Proc. IEE, Third International Conference on Holographic Systems, Components and Applications, **342**, 57-61, Edinburgh, U.K. (1991).

[BLA90] L. T. Blair and L. Solymar, "Grating profiles in dichromated gelatin.", Optics Communications, **77**, 365-366 (1990).

[BLA91] L. T. Blair and L. Solymar, "Double exposure planar transmission holograms recorded in nonlinear dichromated gelatin.", Applied Optics, **30**, 775-779 (1991).

[BLA95] P. Blair, M. R. Taghizadeh, W. Parkes and C. D. W. Wilkinson, "High-efficiency binary fan-out gratings by modulation of a high-frequency carrier grating.", Applied Optics, **34**, 2406-2413 (1995).

[BOR91] N. F. Borelli, "The generation of lens arrays using photothermal techniques.", Proc. IOP, Microlens Arrays, Short Meetings Series, **30**, 1-16, Teddington, U.K. (1991).

[BOT81] L. C. Botten, M. S. Craig, R. C. McPhedran, J. L. Adams and J. R. Andrewartha, "The dielectric lamellar diffraction grating.", Optica Acta, **28**, 413-428 (1981).

[BOY83] R. W. Boyd, "Radiometry and the detection of optical radiation.", J. Wiley & Sons, New York, U.S.A. (1983).

[BRE88] K. -H. Brenner and A. Huang, "Optical implementation of the perfect shuffle interconnection.", Applied Optics, **27**, 135-137 (1988).

[BUL89] G. S. Buller, "Properties and potential applications of nonlinear interference filters.", Ph.D. thesis, Heriot-Watt University, U.K. (1989).

[BUR66] C. B. Burckhardt, "Diffraction of a plane wave at a sinusoidally stratified dielectric grating.", J. Opt. Soc. Am., **56**, 1502-1509 (1966).

[CAS75] S. K. Case, "Coupled-wave theory for multiply exposed thick holographic gratings.", J. Opt. Soc. Am., **65**, 724-729 (1975).

[CAS76] S. K. Case and R. Alferness, "Index modulation and spatial harmonic generation in DCG films.", Applied Physics, **10**, 41-51 (1976).

[CAS81] S. K. Case, P. R. Haugen and O. J. Lokberg, "Multifacet holographic optical elements for wavefront transformations.", Applied Optics, **20**, 2670-2675 (1981).

[CAW91] P. S. Cawte, "The diffraction efficiency of double-grating holographic fan-out elements.", J. Mod. Opt., **38**, 1365-1376 (1991).

[CHA67] E. B. Champagne, "Non-paraxial imaging, magnification and aberration properties in holography.", J. Opt. Soc. Am., **57**, 51-55 (1967).

[CHA71] M. Chang, "Dichromated gelatin of improved optical quality.", Applied Optics, **10**, 2550-2251 (1971).

[CHA76] B. J. Chang, "Post-processing of developed DCG holograms.", Optics Communications, **17**, 270-272 (1976).

[CHA79] B. J. Chang and C. D. Leonard, "Dichromated gelatin for the fabrication of holographic optical elements.", Applied Optics, **18**, 2407-2417 (1979).

[CHA80] B. J. Chang, "Dichromated gelatin holograms and their applications.", Optical Engineering, **19**, 642-648 (1980).

[CHA93] N. Chateau, "Modelisation et characterisation de reseaux holographiques; etude de composants optiques holographiques en gelatine bichromatee.", Ph.D. thesis (in French), Universite de Paris-Sud centre d'Orsay, France (1993).

[CHA94] N. Chateau and J. P. Hugonin, "Algorithm for the rigorous coupled-wave analysis of grating diffraction.", J. Opt. Soc. Am. A, **11**, 1321-1331 (1994).

[CHE87] H. Chen, R. R. Hershey and E. N. Leith, "Design of a holographic lens for the infra-red.", Applied Optics, **26**, 1983-1988 (1987).

[CHU77a] R. S. Chu and J. A. Kong, "Modal theory of spatially periodic media.", IEEE Trans. Microwave Theory Tech., **MTT-25**, 18-24 (1977).

[CHU77b] R. S. Chu, J. A. Kong and T. Tamir, "Diffraction of Gaussian beams by a periodically modulated layer.", J. Opt. Soc. Am., **67**, 1555-1561 (1977).

[CLO75] D. H. Close, "Holographic optical elements.", Optical Engineering, **14**, 408-419 (1975).

[CLO92] T. J. Cloonan, G. W. Richards, A. L. Lentine, F. B. McCormick and J. R. Erickson, "Free-space photonic switching architectures based on

extended generalised shuffle networks.", Applied Optics, **31**, 7471-7492 (1992).

[CLO93] T. J. Cloonan, "A high bit-rate packet switch architecture with advanced electronic packaging and free-space optical interconnects.", Ph.D. thesis, Heriot-Watt University, U.K. (1993).

[CLO94] T. J. Cloonan, A. L. Lentine, J. M. Sasian, F. B. McCormick, S. J. Hinterlong and H. S. Hinton, "Shuffle-equivalent interconnection topologies based on computer generated binary-phase gratings.", Applied Optics, **33**, 1405-1430 (1994).

[COL94] W. Collinge, "Recovering from M.E. - a guide to self-empowerment.", Souvenir press, London, U.K. (1994).

[CRA91] R. G. A. Craig, "Implementation, control and programming of digital optical systems.", Ph.D. thesis, Heriot-Watt University, U.K. (1991).

[CRA92] N. C. Craft and A. Y. Feldblum, "Optical interconnects based on arrays of surface-emitting lasers and lenslets.", Applied Optics, **31**, 1735-1739 (1992).

[CRO93] W. A. Crossland, R. J. Mears, S. T. Warr and R. W. A. Scar, "Spatial-light-modulator based routing switches.", IOP conference series, **139**, 177-182 (1993).

[CUL82] R. A. Cullen, "Some characteristics of, and measurements on DCG reflection holograms.", Proc. SPIE, **369**, 647-654 (1982).

[CUT66] L. J. Cutrona, A. N. Leith, L. J. Porcello and W. E. Vivian, "On the application of coherent optical processing techniques to synthetic aperture radar.", Proc. IEEE, **54**, 1026-1032 (1966).

[DAM71] H. Dammann and K. Gortler, "High efficiency in-line multiple imaging by means of multiple phase holograms.", Optics Communications, **3**, 312-315 (1971).

[DAV92] N. Davidson, A. A. Friesem and E. Hasman, "Realization of perfect shuffle and inverse perfect shuffle transforms with holographic elements.", Applied Optics, **31**, 1810-1812 (1992).

[DEN80] J. B. Dennis, "Data-flow supercomputers.", Computer, **13**, 48-56, November (1980).

[DER89] M. W. Derstine and A. Guha, "Design considerations for an optical symbolic processing architecture.", Optical Engineering, **28**, 434-446 (1989).

[DES95] M. P. Y. Desmulliez, F. A. P. Tooley, J. A. B. Dines, N. L. Grant, D. J. Goodwill, D. Baillie, B. S. Wherrett, P. W. Foulk, S. Ashcroft and P. Black, "Perfect-shuffle interconnected bitonic sorter: optoelectronic design.", Applied Optics, **34**, 5077-5090 (1995).

[DRU59] P. Drude, "The theory of optics.", Dover Publications, New York, U.S.A. (1959).

[DUF80] W. J. Duffin, "Electricity and magnetism.", McGraw-Hill, London, U.K. (1980).

[DUF86] M. J. B. Duff and T. J. Fountain, "Cellular logic image processing.", Academic Press, San Diego, U.S.A. (1986).

[DUN94] J. Dunlop and D. G. Smith, "Telecommunications engineering.", Chapman & Hall, London, U.K. (1994).

[EVA89] R. W. Evans, A. P. Ramsbottom and D. W. Sheel, "Head-up displays in motor cars.", Proc. IEE, Second International Conference on Holographic Systems, Components and Applications, **311**, 56-57, Bath, U.K. (1989).

[FAL91] O. Falkenstorfer, H. Kobolla, U. Krackhardt, N. Lindlein, J. Schwider, N. Streibl, R. Volkel and H. Weissmann, "Optimization of holographic lenslets and their measurement.", Proc. IOP, Microlens Arrays, Short Meetings Series, **30**, 53-60, Teddington, U.K. (1991).

[FEL88] M. R. Feldman, S. C. Esener, C. C. Guest and S. H. Lee, "Comparison between optical and electrical interconnects based on power and speed considerations.", Applied Optics, **27**, 1742-1751 (1988).

[FEL89] M. R. Feldman and C. C. Guest, "Interconnect density capabilities of computer generated holograms for optical interconnection of very large scale integration circuits.", Applied Optics, **28**, 3134-3137 (1989).

[FOX96] J. Fox, "Surviving M.E. - practical strategies for coming to terms with chronic fatigue.", Vermilion (Ebury Press), Random House, London, U.K. (1996).

[FRI86] A. T. Friberg, "Radiance theorem with partially coherent light.", Optica Acta, **33**, 1369-1376 (1986).

[GAB48] D. Gabor, "A new microscopic principle.", Nature, **161**, 777-778 (1948).

[GAY85] T. K. Gaylord and M. G. Moharam, "Analysis and applications of optical diffraction by gratings.", Proc. IEEE, **73**, 894-937 (1985).

[GEO87] T. G. Georgekutty and H. -K. Liu, "Simplified dichromated gelatin hologram recording process.", Applied Optics, **26**, 372-376 (1987).

[GLU93] E. Gluch, H. Kobolla, K. Zurl, N. Streibland J. Schwider, "Demonstration for an optoelectronic switching network.", J. Mod. Opt, **40**, 1857-1869 (1993).

[GLY87] E. N. Glytsis and T. K. Gaylord, "Rigorous three-dimensional coupled-wave diffraction analysis of single and cascaded anisotropic gratings.", J. Opt. Soc. Am. A, **4**, 2061-2080 (1987).

[GLY89] E. N. Glytsis and T. K. Gaylord, "Rigorous 3-D coupled wave diffraction analysis of multiple superposed gratings in anisotropic media.", Applied Optics, **28**, 2401-2421 (1989).

[GOO78] J. W. Goodman, A. R. Dias and L. M. Woody, "Fully parallel, high speed incoherent optical method for performing discrete Fourier transforms.", Optics Letters, **2**, 1-3 (1978).

[GOO85] J. W. Goodman, "Fan-in and fan-out with optical interconnects.", Optica Acta, **32**, 1489-1496 (1985).

[GRA78] A. Graube, "Holographic optical element materials research.", Report for the U.S. Air Force Office of Scientific Research, Hughes Research Laboratories, U.S.A. (1978).

[GUR85] J. R. Gurd, C. C.Kirkham and I. Watson, "The Manchester prototype dataflow computer.", Communications of the ACM, **28**, 34-52, January (1985).

[HAL89] D. R. Hall and P. E. Jackson, *editors*, "The physics and technology of laser resonators.", Adam Hilger, Bristol, U.K. (1989).

[HAL93] T. J. Hall, L. Q. Xu, S. Jamieson, P. Dainty, K. B. Russ, S. R. Cherry, N. C. Roberts and A. L. Walton, "Free-space optical networks using diffractive and refractive micro-optics.", Proc. IEE, Fourth International Conference on Holographic Systems, Components and Applications, **379**, 35-40, Neuchatel, Switzerland (1993).

[HAR84] P. Hariharan, "Optical holography - principles, techniques and applications.", Cambridge University press, Cambridge, U.K. (1984).

[HAR92] J. E. Harvey and E. A. Nevis, "Angular grating anomalies - effects of finite beam size on wide-angle diffraction phenomena.", Applied Optics, **31**, 6782-6788 (1992).

[HAU89] H. J. Haumann, J. Hausmann, H. Kobolla, F. Sauer, W. Stork, N. Streibl, M. Testorf and R. Volkel, "Holographic optical interconnects in dichromated gelatin (DCG).", Proc. International Symposium on Optics in Computing, 137-142, Toulouse, France (1989).

[HAU91] H. J. Haumann, H. Kobolla, F. Sauer, J. Schmidt, J. Schwider, W. Stork, N. Streibl and R. Volkel, "Optoelectronic interconnection based on a light-guiding plate with holographic coupling elements.", Optical Engineering, **30**, 1620-1623 (1991).

[HEC87] E. Hecht, "Optics.", Second Edition, Addison-Wesley Publishing Co., Reading Ma, U.S.A. (1987).

[HER86] H. P. Herzig, "Holographic optical elements (HOE) for semiconductor lasers.", Optics Communications, **58**, 144-148 (1986).

[HER92] H. P. Herzig, P. Ehbets, D. Prongue and R. Dandliker, "Fan-out elements recorded as volume holograms: optimized recording conditions.", Applied Optics, **31**, 5716-5723 (1992).

[HER93] H. P. Herzig, P. Ehbets and R. Dandliker, "Potential of holographically recorded fan-out elements.", Proc. IEE, Fourth International Conference on Holographic Systems, Components and Applications, **379**, 198-203, Neuchatel, Switzerland (1993).

[HIL85] W. D. Hillis, "The connection machine.", MIT press, Cambridge Ma, U.S.A. (1985).

[HIN94] H. S. Hinton, T. J. Cloonan, F. B. McCormick, A. L. Lentine and F. A. P. Tooley, "Free-space digital optical-systems.", Proc. IEEE, **82**, 1632-1649 (1994).

[HUA84] A. Huang, "Architectural considerations involved in the design of an optical digital computer.", Proc. IEEE, **72**, 780-786 (1984).

[HUT91] M. C. Hutley, D. Daly and R. F. Stevens, "The testing of microlens arrays.", Proc. IOP, Microlens Arrays, Short Meetings Series, **30**, 67-81, Teddington, U.K. (1991).

[HUT92] M. C. Hutley, P. Savander and M. Schrader, "The use of microlenses for making spatially variant optical interconnects.", Pure Applied Optics,

1, 337-346 (1992).

[ICH92] H. Ichikawa, "Optical beam array generation with phase gratings.", Ph.D. thesis, Heriot-Watt University, U.K. (1992).

[ICH93] H. Ichikawa, J. Turunen and M. R. Taghizadeh, "Analysis of hybrid holographic gratings by thin-grating decomposition method.", J. Opt. Soc. Am. A, **10**, 1176-1183 (1993).

[ING89] R. T. Ingwall and M. Troll, "Mechanism of hologram formation in DMP-128 photopolymer.", Optical Engineering, **26**, 586-591 (1989).

[JAA92] T. Jaaskelainen and M. Kuittinen, "Planar interconnection gratings.", Optical Computing and Processing, **2**, 29-38 (1992).

[JAC95] J. L. Jackel and W. J. Tomlinson, "An integrated optical illusion: the lossless single-mode waveguide combiner.", Optics & Photonics News, O.S.A., **6**, n10, 47-50 (1995).

[JAH88] J. Jahns and M. J. Murdocca, "Crossover networks and their optical implementation.", Applied Optics, **27**, 3155-3160 (1988).

[JAH90a] J. Jahns and B. A. Brumback, "Integrated optical split-and-shift module based on planar optics.", Optics Communications, **76**, 318-320 (1990).

[JAH90b] J. Jahns and W. Daschner, "Optical cyclic shifter using diffractive lenslet arrays.", Optics Communications, **79**, 407-410 (1990).

[JAH92] J. Jahns, Y. H. Lee, C. A. Burrus and J. L. Jewell, "Optical interconnects using top-surface-emitting microlasers and planar optics.", Applied Optics, **31**, 592-597 (1992).

[JAH94a] J. Jahns, F. Sauer, B. Tell, K. F. Brown-Goebler, A. Y. Feldblum, C. R. Nijander and W. P. Townsend, "Parallel optical interconnections using surface-emitting microlasers and a hybrid imaging system.", Optics Communications, **109**, 328-337 (1994).

[JAH94b] J. Jahns, "Planar packaging of free-space optical interconnections.", Proc. IEEE, **82**, 1623-1631 (1994).

[JEN84] B. K. Jenkins, P. Chavel, R. Forcheimer, A. A. Sawchuk and T. C. Strand, "Architectural implications of a digital optical processor.", Applied Optics, **23**, 3465-3474 (1984).

[KAS73] F. G. Kaspar, "Diffraction by thick, periodically stratified gratings with complex dielectric constant.", J. Opt. Soc. Am., **63**, 37-45 (1973).

[KEI93] T. Keininen and O. Salminen, "Dichromated gelatin reflection gratings, influence of secondary gratings.", J. Mod. Opt., **40**, 1149-1454 (1993).

[KIR91] A. G. Kirk, W. A. Crossland and T. J. Hall, "A compact and scalable free-space optical crossbar.", Proc. IEE, Third International Conference on Holographic Systems, Components and Applications, **342**, 137-141, Edinburgh, U.K. (1991).

[KLE67] W. R. Klein and B. D. Cook, "Unified approach to ultrasonic light diffraction.", IEEE Trans. Sonics Ultrason., **SU-14**, 123-134 (1967).

[KNO78] K. Knop, "Rigorous diffraction theory for transmission phase gratings with deep rectangular grooves.", J. Opt. Soc. Am., **68**, 1206-1210 (1978).

[KOB93] H. Kobolla, J. T. Sheridan, E. Gluch, J. Schmidt, R. Volkel, J. Schwider and N. Streibl, "Holographic 2D polarization deflection elements.", J. Mod. Opt., **40**, 613-624 (1993).

[KOB94] H. Kobolla, "Dichromated gelatine transmission holograms for 1.5μm wavelength.", J. Mod. Opt., **41**, 19-29 (1994).

[KOB95] H. Kobolla, J. Schmidt, E. Gluch and J. Schwider, "Holographic perfect shuffle permutation element for a miniaturized switching network.", Applied Optics, **34**, 2844-2847 (1995).

[KOG69] H. Kogelnik, "Coupled wave theory for thick hologram gratings.", The Bell System Technical Journal, **48**, 2909-2947 (1969).

[KON77] J. A. Kong, "Second-order coupled-mode equations for spatially periodic media.", J. Opt. Soc. Am., **67**, 1165-1170 (1977).

[KOS85] R. K. Kostuk, J. W. Goodman and L. Hesselink, "Optical imaging applied to microelectronic chip-to-chip interconnections.", Applied Optics, **24**, 2851-2858 (1985).

[KOS86] R. K. Kostuk, J. W. Goodman and L. Hesselink, "Volume reflection holograms with multiple gratings: an experimental and theoretical evaluation.", Applied Optics, **25**, 4362-4369 (1986).

[KOS89a] R. K. Kostuk, "Comparison of models for multiplexed holograms.", Applied Optics, **28**, 771-777 (1989).

[KOS89b] R. K. Kostuk, Y-T. Huang, D. Hetherington and M. Kato, "Reducing alignment and chromatic sensitivity of holographic optical interconnects with substrate-mode holograms.", Applied Optics, **28**, 4939-4944 (1989).

[KOS93] R. K. Kostuk, J.-H. Yeh and M. Fink, "Distributed optical data bus for board-level interconnects.", Applied Optics, **32**, 5010-5021 (1993).

[KOW76] R. Kowarschik, "Diffraction efficiency of attenuated sinusoidally modulated gratings in volume holograms.", Optica Acta, **23**, 1039-1051 (1976).

[KOW78a] R. Kowarschik, "Diffraction efficiency of sequentially stored gratings in transmission volume holograms.", Optica Acta, **25**, 67-81 (1978).

[KOW78b] R. Kowarschik, "Diffraction efficiency of sequentially stored gratings in reflection volume holograms.", Optical and Quantum Electronics, **10**, 171-178 (1978).

[LAM68] J. T. LaMacchia and C. J. Vincelette, "Comparison of the diffraction efficiency of multiple exposure and single exposure holograms.", Applied Optics, **7**, 1857-1858 (1968).

[LAM86] L. Lamport, "LaTeX: A document preparation system.", Addison-Wesley Publishing Co., Reading Ma, U.S.A. (1986).

[LAT71a] J. N. Latta, "Computer-based analysis of hologram imagery and aberrations 2: Aberrations induced by a wavelength shift.", Applied Optics, **10**, 609-618 (1971).

[LAT71b] J. N. Latta, "Computer-based analysis of holography using ray-tracing.", Applied Optics, **10**, 2698-2710 (1971).

[LAT72] J. N. Latta, "Analysis of multiple optical elements with low dispersion and low aberrations.", Applied Optics, **11**, 1686-1696 (1972).

[LAT79] M. R. Latta and R. V. Pole, "Design techniques for forming 488nm holographic lenses with reconstruction at 633nm.", Applied Optics, **18**, 2418-2421 (1979).

[LED80] F. Lederer and U. Langbein, "Modal theory for thick holographic gratings with sharp boundaries II. Unslanted transmission and reflection gratings.", Optica Acta, **27**, 183-200 (1980).

[LEE87] S. H. Lee, "Computer generated holography: an introduction.", Applied Optics, **26**, 4350 (1987).

[LEG87] J. R. Leger, G. J. Swanson and W. B. Veldkamp, "Coherent laser addition using binary phase gratings.", Applied Optics, **26**, 4391-4399 (1987).

[LEI62] E. N. Leith and J. Upatnieks, "Reconstructed wavefronts and communication theory.", J. Opt. Soc. Am., **52**, 1123-1130 (1962).

[LEI65] E. N. Leith, J. Upatnieks and K. A. Haines, "Microscopy by wavefront reconstruction.", J. Opt. Soc. Am., **55**, 981-986 (1965).

[LEN93] A. L. Lentine, "Issues in the design of self electro-optic effect devices for optical signal processing.", Ph.D. thesis, Heriot-Watt University, U.K. (1993).

[LEN96] A. L. Lentine, K. W. Goossen, J. A. Walker, L. M. F. Chirovsky, L. A. D'Asaro, S. P. Hui, B. T. Tseng, R. E. Leibenguth, D. P. Kossives, D. W. Dahringer, D. D. Bacon, T. K. Woodward and D. A. B. Miller, "Arrays of optoelectronic switching nodes comprised of flip-chip-bonded MQW modulators and detectors on silicon CMOS circuitry.", IEEE Photonics Technology Letters, **8**, 221-223 (1996).

[LES69] L. B. Lesem, P. M. Hirsch and J. A. Jordan, "The Kinoform: A new wavefront reconstruction device.", IBM J. Res. Develop., **13**, 150-155 (1969).

[LEW83] J. W. Lewis and L. Solymar, "Spurious waves in thick phase gratings.", Optics Communications, **47**, 23-26 (1983).

[LI93a] L. Li and C. W. Haggans, "Convergence of the coupled-wave method for metallic lamellar diffraction gratings.", J. Opt. Soc. Am. A, **10**, 1184-1189 (1993).

[LI93b] L. Li, "Multilayer modal method for diffraction gratings of arbitrary profile, depth, and permittivity.", J. Opt. Soc. Am. A, **10**, 2581-2591 (1993).

[LI94] L. Li, "Multilayer modal method for diffraction gratings of arbitrary profile, depth, and permittivity: addendum.", J. Opt. Soc. Am. A, **11**, 1685 (1994).

[LIN71] L. H. Lin and E. T. Doherty, "Efficient and aberration-free wavefront reconstruction from holograms illuminated at wavelengths differing from the forming wavelength.", Applied Optics, **10**, 1314-1318 (1971).

[LIN92] L. Lindvold, "DCG as a holographic medium.", Ph.D. thesis, Phillips research, Netherlands (1992).

[LOH86] A. W. Lohmann, "What classical optics can do for the digital optical computer.", Applied Optics, **25**, 1543-1549 (1986).

[LUD73] U. W. Ludwig, "Generalised grating ray tracing equations.", J. Opt. Soc. Am., **63**, 1105-1107 (1973).

[MAG77] R. Magnusson and T. K. Gaylord, "Analysis of multiwave diffraction of thick gratings.", J. Opt. Soc. Am., **67**, 1165-1170 (1977).

[MAG78a] R. Magnusson and T. K. Gaylord, "Diffraction efficiencies of thin phase gratings with grating shape.", J. Opt. Soc. Am., **68**, 806-809 (1978).

[MAG78b] R. Magnusson and T. K. Gaylord, "Equivalence of multiwave coupled-wave theory and modal theory for periodic-media diffraction.", J. Opt. Soc. Am., **68**, 1777-1779 (1978).

[MAG80] G. A. Mago', "A cellular computer architecture for functional programming.", Proc. IEEE COMPCON 80, 179-187 (1980).

[MAR72] D. Marcuse, "Light transmission optics.", Van Nostrand-Reinhold, Princeton, U.S.A. (1972).

[MAY72] D. Maystre and R. Petit, "Diffraction par un reseau lamellaire infinement conducteur.", Optics Communications, **5**, 90-93 (1972).

[MAY80] D. Maystre, *Integral methods* chapter 3 in, "Electromagnetic theory of gratings.", Topics in current physics, **22**, *editor*: R. Petit, Springer-Verlag, Berlin (1980).

[MAY84] D. Maystre, "Rigorous vector theories of diffraction gratings.", in Progress in Optics **XXI**, 1-67, *editor*: E. Wolf, North-Holland Physics Publishing, (1984).

[MCA86] A. D. McAulay, "Optical crossbar interconnected digital signal processor with basic algorithms.", Optical Engineering, **25**, 82-90 (1986).

[MCA91] A. D. McAulay, "Optical computer architectures.", J. Wiley & Sons, New York, U.S.A. (1991).

[MCC90] F. B. McCormick and M. E. Prise, "Optical circuitry for free-space interconnections.", Applied Optics, **29**, 2013-2018 (1990).

[MCC91] F. B. McCormick, F. A. P. Tooley, T. J. Cloonan, J. M. Sasian and H. S. Hinton, "Microbeam optical interconnections using microlens arrays.", OSA Proceedings on Photonic Switching, *editors*: H. S. Hinton & J. W. Goodman, **8**, 90-96 (1991).

[MCC93] F. B. McCormick, "Optical and mechanical design issues in free-space optical interconnection and processing.", Ph.D. thesis, Heriot-Watt University, U.K. (1993).

[MCG80] S. P. McGrew, "Colour control in DCG reflection holograms.", Proc. SPIE, **215**, 24 (1980).

[MEA80] C. Mead and L. Conway, "Introduction to VLSI systems.", Addison-Wesley, Reading Ma, U.S.A. (1980).

[MEY77] D. Meyerhofer, *Dichromated gelatin* chapter 3 in "Holographic recording materials.", Topics in Applied Physics, **20**, *editor*: H. M. Smith, Springer-Verlag, Berlin, Germany (1977).

[MIL84] D. A. B. Miller, D. S. Chemla, T. C. Damen, A. C. Gossard, W. Wiegmann, T. H. Wood and C. A. Burns, "Novel hybrid optically bistable switch: the quantum well self-electro-optic effect device.", Applied Physics Letters, **45**, 13-15 (1984).

[MIL93] J. M. Miller, "Synthesis of high efficiency diffractive optical elements.", Ph.D. thesis, Heriot-Watt University, U.K. (1993).

[MIL94] J. M. Miller, J. Turunen, E. Noponen, A. Vasara and M. R. Taghizadeh, "Rigorous modal theory for multiply grooved lamellar gratings.", Optics Communications, **111**, 526-535 (1994).

[MIL95] D. A. B. Miller, "Hybrid SEED - massively parallel optical interconnections for silicon ICs.", Proc. IEEE, Second International Conference on

Massively Parallel Processing using Optical Interconnections, 2-7, San Antonio, U.S.A. (1995).

[MOH80a] M. G. Moharam, T. K. Gaylord and R. Magnusson, "Criteria for Bragg regime diffraction by phase gratings.", Optics Communications, **32**, 14-18 (1980).

[MOH80b] M. G. Moharam, T. K. Gaylord and R. Magnusson, "Bragg diffraction of finite beams by thick gratings.", J. Opt. Soc. Am., **70**, 300-304 (1980).

[MOH81a] M. G. Moharam and T. K. Gaylord, "Coupled-wave analysis of reflection gratings.", Applied Optics, **20**, 240-244 (1981).

[MOH81b] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of planar-grating diffraction.", J. Opt. Soc. Am., **71**, 811-818 (1981).

[MOH82a] M. G. Moharam and T. K. Gaylord, "Chain-matrix analysis of arbitrary-thickness dielectric reflection gratings.", J. Opt. Soc. Am., **72**, 187-190 (1982).

[MOH82b] M. G. Moharam and T. K. Gaylord, "Diffraction analysis of dielectric surface-relief gratings.", J. Opt. Soc. Am., **72**, 1385-1392 (1982).

[MOH83a] M. G. Moharam and T. K. Gaylord, "Comments on analyses of reflection gratings.", J. Opt. Soc. Am., **73**, 399-401 (1983).

[MOH83b] M. G. Moharam and T. K. Gaylord, "Rigorous coupled-wave analysis of grating diffraction - E–mode polarisation and losses.", J. Opt. Soc. Am., **73**, 451-455 (1983).

[MOH83c] M. G. Moharam and T. K. Gaylord, "Three-dimensional vector coupled-wave analysis of planar grating diffraction.", J. Opt. Soc. Am., **73**, 1105-1112 (1983).

[MOH95a] M. G. Moharam, E. B. Grann, D. A. Pommet and T. K. Gaylord, "Formulation for stable and efficient implementation of the rigorous coupled-

wave analysis of binary gratings.", J. Opt. Soc. Am., **12**, 1068-1076 (1995).

[MOH95b] M. G. Moharam, D. A. Pommet, E. B. Grann and T. K. Gaylord, "Stable implementation of the rigorous coupled-wave analysis for surface-relief gratings: enhanced transmittance matrix approach.", J. Opt. Soc. Am., **12**, 1077-1086 (1995).

[MON93] D. M. Monro, J. A. Dallas, J. A. Nicholls, M. D. Cripps and W. A. Crossland, "An optically connected parallel machine.", Proc. OSA, Photonic Switching, Technical Digest Series, **16**, 280-283 (1993).

[MUR88] M. J. Murdocca, "Theory and applications of free-space digital optical computing.", Ph.D. thesis, Rutgers University, U.S.A. (1988).

[NAG93] The Numerical Algorithms Group Limited, The NAG Fortran library, Mark 16, Oxford, U.K. (1993).

[NAI90] G. M. Naik, A. Mathur and S. V. Pappu, "Dichromated gelatin holograms: an investigation into their environmental stability.", Applied Optics, **29**, 5292-5297 (1990).

[NEW87] J. C. W. Newell, "Optical holography in dichromated gelatin.", Ph.D. thesis, Oxford University, U.K. (1987).

[NOP92] E. Noponen, "Rigorous diffraction theory of computer-generated gratings.", Licentiate thesis, Helsinki University of Technology, Finland (1992).

[NOP93] E. Noponen, J. Turunen and A. Vasara, "Electromagnetic theory and the design of diffractive lens arrays.", J. Opt. Soc. Am. A, **10**, 434-443 (1993).

[NOP94a] E. Noponen, "Electromagnetic theory of diffractive optics.", Ph.D. thesis, Helsinki University of Technology, Finland (1994).

[NOP94b] E. Noponen and J. Turunen, "Binary high-frequency-carrier diffractive optical elements: electromagnetic theory.", J. Opt. Soc. Am. A, **11**, 1097-1109 (1994).

[OLI84] J. Oliva, P. G. Boj, and M. Pardo, "Dichromated gelatin holograms derived from Agfa 8E75HD plates.", Applied Optics, **23**, 196-197 (1984).

[ORA94] Optical Research Associates Limited, Code V ray-tracing, version 8.03, Pasadena Ca, U.S.A. (1994).

[PAS92] I. Pascual, A. Belendez and A. Fimia, "Analysis of the holographic reciprocity law for dichromated gelatin.", Applied Optics, **31**, 3200-3201 (1992).

[PEN95] S. Peng and G. M. Morris, "Efficient implementation of rigorous coupled-wave analysis for surface-relief gratings.", J. Opt. Soc. Am. A, **12**, 1087-1096 (1995).

[PET80] R. Petit, *editor*, "Electromagnetic theory of gratings.", Topics in current physics, **22**, Springer-Verlag, Berlin, Germany (1980).

[PHA56] P. Phariseau, "On the diffraction of light by progressive supersonic waves.", Proc. Ind. Acad. Sci. A, **44**, 165-170 (1956).

[PLA95] D. V. Plant, B. Robertson, H. S. Hinto, W. M. Robertson, G. C. Boisset, N. H. Kim, Y. S. Liu, M. R. Otazo, D. R. Rolston and A. Z.Shang, "An optical backplane demonstrator system based on FET-SEED smart pixel arrays and diffractive lenslet arrays.", IEEE Photonics Technology Letters, **7**, 1057-1059 (1995).

[PRI88] M. E. Prise, N. Streibl, and M. M. Downs, "Optical considerations in the design of digital optical computers.", Optical and Quantum Electronics, **20**, 49-77 (1988).

[PRI90] M. E. Prise, N. C. Craft, R. E. LaMarche, M. M. Downs, S. J. Walker, L. A. D'Asaro and L. M. F. Chirovsky, "A module for optical logic cir-

cuits using symmetric self electro-optic effect devices.", Applied Optics, **29**, 2164-2170 (1990).

[PRI95] S. M. Prince, F. A. P. Tooley, S. Wakelin and M. R. Taghizadeh, "Implementation of optical perfect-shuffle module.", Applied Optics, **34**, 1775-1782 (1995).

[PRO89] D. Prongue and H. P. Herzig, "Design and fabrication of HOE for clock distribution in integrated circuits.", Proc. IEE, Second International Conference on Holographic Systems, Components and Applications, **311**, 204-208, Bath, U.K. (1989).

[PSA84] D. Psaltis, "Two-dimensional optical processing using one-dimensional input devices.", Proc. IEEE, **72**, 962-974 (1984).

[RAM35] C. V. Raman and N. S. N. Nath, "The diffraction of light by high frequency sound waves.", *Parts I & II*, Proc. Ind. Acad. Sci. A, **2**, 406-420 (1935).

[RAM36] C. V. Raman and N. S. N. Nath, "The diffraction of light by high frequency sound waves.", *Parts III - V*, Proc. Ind. Acad. Sci. A, **3**, 75-84, 119-125, 459-465 (1936).

[RAY07] Lord Rayleigh, "On the dynamical theory of gratings.", Proc. Royal Society, **A79**, 399-416 (1907).

[REB89] J. M. Rebordao and A. A. Andrade, "Microwave drying effects on dichromated gelatin holograms.", Applied Optics, **28**, 4393-4400 (1989).

[RED88] I. R. Redmond and M. R. Taghizadeh, "Continuously variable laser beam attenuator.", European patent no. 88109667.1 (1988).

[RED89a] I. R. Redmond, "Holographic optical elements in dichromated gelatin.", Ph.D. thesis, Heriot-Watt University, U.K. (1989).

[RED89b] I. R. Redmond, E. J. Restall and A. C. Walker, "High performance holographic optics for the visible and near infra-red.", Proc. IEE, Second International Conference on Holographic Systems, Components and Applications, **311**, 190-194, Bath, U.K. (1989).

[RES89] E. J. Restall, "Feasible materials for long-term optical memories.", Boeing project report, Physics Department, Heriot-Watt University, U.K., May (1989).

[RES91a] E. J. Restall, B. Robertson, M. R. Taghizadeh and A. C. Walker, "Two dimensional spatially variant optical interconnects.", Proc. OSA, Optical Computing, Technical Digest Series, **6**, 49-52, Salt Lake City, U.S.A. (1991).

[RES91b] E. J. Restall, B. Robertson, M. R. Taghizadeh and A. C. Walker, "Holographic microlens arrays as spatially variant optical interconnects.", Proc. IOP, Microlens Arrays, Short Meetings Series, **30**, 83-89, Teddington, U.K. (1991).

[RES91c] E. J. Restall, I. R. Redmond, and A. C. Walker, "A nearest neighbour spatially invariant reflection interconnect in dichromated gelatin.", Proc. IEE, Third International Conference on Holographic Systems, Components and Applications, **342**, 40-44, Edinburgh, U.K. (1991).

[RES91d] E. J. Restall, B. Robertson, M. R. Taghizadeh, and A. C. Walker, "Two dimensional non-local multi-facet holographic interconnects in dichromated gelatin.", Proc. IEE, Third International Conference on Holographic Systems, Components and Applications, **342**, 127-131, Edinburgh, U.K. (1991).

[ROB88] N. C. Roberts, "Free space holographic optical interconnects.", Proc. IEE, Optical Interconnects, Technical Digest Series, **121**, $^6/_1$ - $^6/_4$ (1988).

[ROB90] B. Robertson, E. J. Restall, M. R. Taghizadeh and A. C. Walker, "Space-variant interconnection networks in dichromated gelatin.", Proc. SPIE,

1319, 128-129, (1990).

[ROB91a] B. Robertson, E. J. Restall, M. R. Taghizadeh and A. C. Walker, "Space-variant holographic optical elements in dichromated gelatin.", Applied Optics, **30**, 2368-2375 (1991).

[ROB91b] B. Robertson, J. Turunen, H. Ichikawa, J. M. Miller, M. R. Taghizadeh and A. Vasara, "Hybrid kinoform fan-out holograms in dichromated gelatin.", Applied Optics, **30**, 3711-3720 (1991).

[ROB93a] B. Robertson, "Free-space holographic optical interconnects in dichromated gelatin.", Ph.D. thesis, Heriot-Watt University, U.K. (1993).

[ROB93b] B. Robertson, C. Godsalve and M. R. Taghizadeh, "DCG holography: An investigation into laser induced damage.", Applied Optics, **32**, 6587-6597 (1993).

[RUS81] P. St. J. Russell, "Optical volume holography.", Physics Reports, **71**, 209-312 (1981).

[RUS84] P. St. J. Russell, "Power conservation and field structures in uniform dielectric gratings.", J. Opt. Soc. Am. A, **1**, 293-299 (1984).

[SAA95] J. Saarinen, "Computer-synthesized diffractive elements in micro-optics.", Ph.D. thesis, Helsinki University of Technology, Finland (1995).

[SAM80] D. M. Samoilovich, A. Zeichner, A. A. Friesem, "The mechanism of volume hologram formation in dichromated gelatin.", Photographic Science and Engineering, **24**, 161 (1980).

[SAU89] F. Sauer, "Fabrication of diffractive-reflective optical interconnects for infrared operation based on total internal reflection.", Applied Optics, **28**, 386-388 (1989).

[SAU94] F. Sauer, J. Jahns, C. S. Nijander, A. V. Feldblum and W. P. Townsend, "Refractive-diffractive micro-optics for permutation interconnects.", Optical Engineering, **33**, 1550-1560 (1994).

[SAW87] A. A. Sawchuk, B. K. Jenkins, C. S. Raghavendra and A. Varma, "Optical crossbar networks.", Computer, **20**, 50-60, June (1987).

[SCH92] J. Schwider, W. Stork, N. Streibl and R. Volkel, "Possibilities and limitations of space-variant holographic optical elements for switching networks and general interconnects.", Applied Optics, **31**, 7403-7410 (1992).

[SED88] R. Sedgewick, "Algorithms.", Addison-Wesley, Reading Ma, U.S.A. (1988).

[SHA68] T. A. Shankoff, "Phase holograms in dichromated gelatin.", Applied Optics, **7**, 2101-2105 (1968).

[SHE82] P. Sheng, R. S. Stepleman and P. N. Sanda, "Exact eigenfunctions for square wave gratings: Application to diffraction and surface-plasmon calculations.", Phys. Rev. B, **26**, 2907-2916 (1982).

[SIE82] D. P. Siewiorek, C. G. Bell and A. Newell, "Computer structures: principles and examples.", McGraw-Hill, London, U.K. (1982).

[SLI85] C. W. Slinger, "Multiple holographic transmission gratings in silver halide emulsion.", Ph.D. thesis, Oxford University, U.K. (1985).

[SLI86] C. W. Slinger and L. Solymar, "Grating interactions in holograms recorded with two object waves.", Applied Optics, **25**, 3283-3287 (1986).

[SLI87] C. W. Slinger, R. R. A. Syms and L. Solymar, "Multiple holographic transmission gratings in silver halide emulsion.", Applied Physics B, **42**, 121-128 (1987).

[SMI77] H. M. Smith, *editor*, "Holographic recording materials.", Topics in applied physics, **20**, Springer-Verlag, Berlin, Germany (1977).

[SMI87] S. D. Smith, A. C. Walker, F. A. P. Tooley and B. S. Wherrett, "The demonstration of restoring digital optical logic.", Nature, **325**, 27-31 (1987).

[SMI95] B. J. Smith, "Interconnection networks for shared memory parallel computers.", Proc. IEEE, Second International Conference on Massively Parallel Processing using Optical Interconnections, 255-256, San Antonio, U.S.A. (1995).

[SMO89] W. K. Smothers, T. J. Trout, A. M. Weber and D. J. Mickish, "Hologram recording in DuPont's new photopolymer materials.", Proc. IEE, Second International Conference on Holographic Systems, Components and Applications, **311**, 184-189, Bath, U.K. (1989).

[SOL77] L. Solymar, "Two-dimensional N-coupled wave theory for volume holograms.", Optics Communications, **23**, 199-202 (1977).

[SOL81] L. Solymar and D. J. Cooke, "Volume holography and volume gratings.", Academic Press, London, U.K. (1981).

[SON93] S. H. Song, C. D. Carey, D. R. Selviah, J. E. Midwinter and E. H. Lee, "Optical perfect-shuffle interconnection using a computer generated hologram.", Applied Optics, **32**, 5022-5025 (1993).

[STI88] C. W. Stirk, R. A. Athale and M. W. Haney, "Folded perfect shuffle optical processor.", Applied Optics, **27**, 202-203 (1988).

[STO71] H. S. Stone, "Parallel processing with the perfect shuffle.", IEEE Trans. Comput., **C-20**, 153-161 (1971).

[STO85] T. Stone and N. George, "Wavelength performance of holographic optical elements.", Applied Optics, **24**, 3797-3810 (1985).

[STO86] T. W. Stone, "Holographic optical elements.", Ph.D. thesis, The Institute of Optics, University of Rochester, U.S.A. (1986).

[STO91] W. Stork, N. Streibl, H. Haidner and P. Kipfer, "Artificial distributed-index media fabricated by zero-order gratings.", Optics Letters, **16**, 1921-1923 (1991).

[STR89] N. Streibl, "Beam shaping with optical array generators.", J. Mod. Opt., **36**, 1559-1573 (1989).

[STR91] N. Streibl, U. Nolscher and J. Jahns, "Array generation with lenslet arrays.", Applied Optics, **30**, 2739-2742 (1991).

[STR93] N. Streibl, R. Volkel, J. Schwider, P. Habel and N. Lindlein, "Parallel optoelectronic interconnections with high packing density through a light-guiding plate using grating couplers and field lenses.", Optics Communications, **99**, 167-171 (1993).

[SU75] S. F. Su and T. K. Gaylord, "Calculation of arbitrary-order diffraction efficiencies of thick gratings of arbitrary grating shape.", J. Opt. Soc. Am., **65**, 59-64 (1975).

[SYM80] R. R. A. Syms and L. Solymar, "A comparison of the optical path and differential equation methods for optically thin phase gratings.", Optical and Quantum Electronics, **12**, 383-391 (1980).

[SYM82] R. R. A. Syms and L. Solymar, "Analysis of volume holographic cylindrical lenses.", J. Opt. Soc. Am., **72**, 179-186 (1982).

[TIB84] J. Tiberghien, "New computer architecture.", Academic Press, San Diego, U.S.A. (1984).

[TOO93] F. A. P. Tooley and S. Wakelin, "Design of a symmetrical-self-electro-optic-effect-device cellular-logic image-processor.", Applied Optics, **32**, 1850-1862 (1993).

[TOO94] F. A. P. Tooley, S. Wakelin and M. R. Taghizadeh, "Interconnects for a symmetrical-self-electro-optic-effect-device cellular-logic image-processor.", Applied Optics, **33**, 1398-1404 (1994).

[TUR88] J. Turunen, A. Vasara, J. Westerholm, G. Jin and A. Salin, "Optimisation and fabrication of grating beam splitters.", J. Phys. D, **21**, S102-S105 (1988).

[TUR94] J. Turunen, personal communication (1994).

[URQ94] K. S. Urquart, P. Marchand, Y. Fainman and S. H. Lee, "Diffractive optics applied to free-space optical interconnects.", Applied Optics, **33**, 3670-3682 (1994).

[VAN64] A. B. Vander Lugt, "Signal detection by complex spatial filtering.", Trans. IEEE Information Theory, **IT-10**, 139-145 (1964).

[VAS89] A. Vasara, "Array generation with periodic Fourier-transform holograms.", Licentiate thesis, Helsinki University of Technology, Finland (1989).

[VAS91] A. Vasara, "Computer-generated gratings.", Ph.D. thesis, Helsinki University of Technology, Finland (1991).

[VAS92] A. Vasara, M. R. Taghizadeh, J. Turunen, J. Westerholm, E. Noponen, H. Ichikawa, J. M. Miller, T. Jaakkola and S. Kuisma, "Binary surface-relief gratings for array illumination in digital optics.", Applied Optics, **31**, 3320-3336 (1992).

[VIN80] P. Vincent, *Differential methods* chapter 4 in, "Electromagnetic theory of gratings.", Topics in current physics, **22**, *editor*: R. Petit, Springer-Verlag, Berlin (1980).

[VIT93] Vitesse Semiconductor Corp., "VSC864A-2 gallium arsenide 64 × 64 crosspoint switch.", Commercial literature, Ca, U.S.A. (1993).

[VOL95] R. Volkel, S. Junger, S. Rosner, J. Schwider, U. Gruhler, H. Schreck, H. Brughardt and K. Wiedeburg, "Optical backplane for a broadband switching system.", Electronics Letters, **31**, 234-235 (1995).

[WAL86] A. C. Walker, "Application of bistable optical logic gate arrays to all-optical digital parallel processing.", Applied Optics, **25**, 1578-1585 (1986).

[WAL90] A. C. Walker, B. S. Wherrett and S. D. Smith, *First implementations of optical digital computing circuits using nonlinear devices* chapter 4 in "Nonlinear photonics.", Springer series in electronics and photonics, **30**, *editors*: H. M. Gibbs, G. Khitrova & N. Peyghambarian, Springer-Verlag, Berlin, Germany (1990).

[WAL91] A. C. Walker, R. G. A. Craig, D. J. McKnight, I. R. Redmond, J. F. Snowdon, G. S. Buller, E. J. Restall, R. A. Wilson, S. Wakelin, N. McArdle, P. Meredith, J. M. Miller, G. MacKinnon, M. R. Taghizadeh, S. D. Smith and B. S. Wherrett, "Design and construction of a programmable optical 16 × 16 array processor.", Proc. OSA, Optical Computing, Technical Digest Series, **6**, 199-202, Salt Lake City, U.S.A. (1991).

[WAL92] S. J. Walker, and J. Jahns, "Optical clock distribution using integrated free-space optics.", Optics Communications, **90**, 359-371 (1992).

[WAN93] J. M. Wang, L. Cheng and A. A. Sawchuk, "Light-efficient two-dimensional perfect shuffles with DuPont photopolymer holograms.", Applied Optics, **32**, 7148-7154 (1993).

[WEL74] W. T. Welford, "Aberrations of the symmetrical optical system.", Academic Press, London, U.K. (1975).

[WEL75] W. T. Welford, "A vector ray tracing equation for hologram lenses of arbitrary shape.", Optics Communications, **14**, 322-323 (1975).

[WEL89] W. T. Welford and R. Winston, "High collection nonimaging optics.", Academic Press, San Diego, U.S.A. (1989).

[WHE90] B. S. Wherrett, R. G. A. Craig, J. F. Snowdon, G. S. Buller, F. A. P. Tooley, S. Bowman, G. S. Pawley, I. R. Redmond, D. J. McKnight, M. R. Taghizadeh, A. C. Walker and S. D. Smith, "Construction and tolerancing of an optical-CLIP.", Proc. SPIE, **1215**, 264-273 (1990).

[WHE95] B. S. Wherrett and P. Chavel, *editors*, "Optical computing.", Institute of Physics Conference Series, **139**, IOP publishing, Bristol, U.K. (1995).

[WHI93] H. J. White, N. A. Brownjohn, C. Stace, G. M. Proudley, A. C. Walker, M. R. Taghizadeh, B. Robertson, C. P. Barrett and M. J. Birch, "Practical demonstration of a freespace optical crossbar switch.", Proc. OSA, Photonic Switching, Technical Digest Series, **16**, 129-132 (1993).

[WIN82] E. Winick, "Designing efficient aberration-free holographic lenses in the presence of a reconstruction wavelength shift.", J. Opt. Soc. Am., **72**, 143-148 (1982).

[WOO83] B. H. Woodcock, "Volume phase holograms and their application to avionic displays.", Proc. SPIE, **399**, 333-338 (1983).

[WYR91] F. Wyrowski and O. Bryngdahl, "Digital holography as part of diffractive optics.", Rep. Prog. Phys., **54**, 1481-1571 (1991).

[ZAL95] D. Zaleta, M. Larsson, W. Daschner and S. H. Lee, "Design methods for space-variant optical interconnections to achieve optimum power throughput.", Applied Optics, **34**, 2436-2447 (1995).

[ZYL83] Z. Zylberberg and E. Marom, "Rigorous coupled-wave analysis of pure reflection gratings.", J. Opt. Soc. Am., **73**, 392-398 (1983).